Spatial Habitat Modeling for a Threatened Plant in a Prairie Sand Dune Landscape

A Thesis Submitted to the College of

Graduate Studies and Research

in Partial Fulfillment of the Requirements

for the Degree of Master of Science

in the Department of Geography and Planning

University of Saskatchewan

Saskatoon

By

Sarah Lowe

Permission to Use

In presenting this thesis in partial fulfillment of the requirements for a Master of Science degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

DISCLAIMER

The software packages mentioned in this thesis were exclusively used to meet the thesis and/or exhibition requirements for the degree of Master of Science at the University of Saskatchewan. Reference in this thesis to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement, recommendation, or favoring by the University of Saskatchewan. The views and opinions of the author expressed herein do not state or reflect those of the University of Saskatchewan, and shall not be used for advertising or product endorsement purposes.

Requests for permission to copy or to make other uses of materials in this thesis in whole or part should be addressed to:

Head of the Department of Geography and Plannning
University of Saskatchewan
Saskatoon, Saskatchewan S7N 5C8
Canada

OR

Dean
College of Graduate Studies and Research
University of Saskatchewan
107 Administration Place
Saskatoon, Saskatchewan S7N 5A2
Canada

Abstract

In 1998, hairy prairie-clover was listed as threatened by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC) and subsequently afforded protection under the Species at Risk Act in 2004. Hairy prairie-clover, being a habitat specialist species confined to areas of sparsely vegetated to bare sand, may provide an indication of the loss of a once viable natural mixed-grass prairie and sand dune landscape indicative to southern Saskatchewan. Therefore, critical habitat identification for hairy prairie-clover is of particular concern, not only to provide conservation efforts for this particular species, but also for bare sand and sand dune environments which are some of the most sensitive landscapes on the prairies. The goal of this thesis is to identify and spatially delineate areas of critical habitat for hairy prairie-clover within the range of a known metapopulation in the Dundurn sand hills south of Saskatoon, Saskatchewan. This research was divided into two specific objectives: 1) to investigate the spatial relationship between bare sand habitat for hairy prairie-clover and other land cover classes, and 2) to study the relationship between habitat configuration and hairy prairie-clover occurrence.

To achieve the first objective, the desired output was a land cover classification of the study site at an appropriate spatial and temporal resolution. Wavelet analysis revealed that the optimum spatial resolution for bare sand identification and delineation in the study site was between 2-5 m. Analysis of field spectroradiometer measurements throughout the growing season concluded that the early and late growing seasons were best for spectrally discriminating between land cover classes. A multi-resolution, multi-temporal land cover classification using object-oriented methods resulted in an overall classification accuracy of 79% with a user's and producer's accuracy of 85% for bare sand. Grassland comprised the matrix of the area covering 45.5% of the study site. Aspen and shrub were the most dominating landscape elements comprising 25.5% and 19.2% of the study site respectively. Bare sand made up only 6.0% of the study site while juniper was the least persistent class comprising only 2.7% of the study site.

The desired output from objective two was a critical habitat landscape mosaic for hairy prairie-clover. Patch scaled metrics were calculated for bare sand patches identified in the land cover classification from objective one. Binary logistic regression was used to identify which metrics could explain and predict hairy prairie-clover occurrences. Results showed that almost 29% of the variation in bare sand patch occupancy could be explained by the size, shape, and degree of isolation of a sand patch as well as the amount of vegetation on a sand patch in the early growing season. Based on these variables, 18.8% of sand patches in the study site were predicted to be unsuitable hairy prairie-clover habitat, 45.7% were predicted to be marginally unsuitable, 32.7% were predicted to be suitable, and 2.8% were predicted to be marginally suitable. Overall prediction accuracy was about 61% with 80% of occurrences and 54% of non-occurrences being correctly predicted.

Acknowledgements

The author wishes to thank Dr. Darcy Henderson with the Canadian Wildlife Service in Saskatoon, SK. and Environment Canada's Interdepartmental Recovery Fund for providing funding and research support. Thank you to Candace Neufeld, Selena Black, Lindsay Hargreaves, Fiona Ross, Monika Tan, Christiane Cattelier, Andrea Fedorchuk, Tessa Giroux, Kimberly Lansdall, Robert Read, Jennalee Mischkolz, and Heather Peat-Hamm for assistance with field surveys and data collection. Thank you to NSERC for providing funding for this project. Thank you to my supervisor Dr. Xulin Guo and committee members Dr. Mohammad Kamal and Dr. Darcy Henderson for their guidance and support throughout this project. Thank you to fellow graduate students Xiaohui Yang, Zhaoqin Li, and Kai Wang for their support.

Table of Contents

List of Figures

List of Tables

1. Introduction and Research Rationale

This chapter will provide an overview of the broader concepts that make up the basis of this study. The rationale for this research will be established within the larger context of the literature. Study objectives and research questions will be established.

1.1 Landscape Pattern Missing in Critical Habitat Identification for Species at Risk

Anthropogenic induced habitat degradation and loss have been identified as the main threats to species survival (Fahrig, 2001; Hodges and Elder, 2008). As a result, the identification of the minimum amount of habitat necessary for species survival is of concern, usually identified in relation to a threshold value in support of a recovery goal and objectives (Fahrig, 2001; Canadian wildlife Service, 2005). Currently, critical habitat designation is most often based on the ecological and biological requirements of a species as determined from a set of environmental variables present in the locations it occupies with only secondary consideration to the spatial distribution of these habitat characteristics (Hoekstra et al., 2002). However, it should be recognized that several factors such as rate of emigration of an organism from a patch, habitat fragmentation, and matrix quality, which have been identified as some of the main factors influencing the determination of threshold values, are influenced by landscape and habitat pattern (Fahrig, 2001). Fahrig and Merriam (1994) specifically identified the need for the incorporation of landscape spatial structure into habitat suitability modelling because it should be considered that all plant populations are spatially structured. Naujokaitis-Lweis et al. (2009) further acknowledges the need for understanding how spatial parameters influence population viability in the long term. Specifically, when dealing with dispersal limited species in fragmented landscapes where small populations are at a high risk of extinction, habitat pattern is most important (Wimberly, 2006). Despite this, information on larger scale spatial patterns of habitat such as space, cover, and distribution are the least likely criteria to be used in critical habitat designation (Hodges and Elder, 2008).

1.2 Pattern and Process in the Landscape

The relationship between landscape pattern and population process is central to conservation ecology because each can affect the other (Franks et al., 2009). A habitat niche is able to capture

the population processes that create a unique spatial pattern and the constraints of that pattern on the population processes (Eriksson, 1996). For example, the dispersal ability of a plant represents its capacity for spatial self-organization, thus affecting the composition and structure of the resulting vegetation community (Alados et al., 2009). However, existing landscape pattern such as topography or soil moisture, can dictate the survival rate of seedlings at a given location, thus reducing individual specie's capacity for self-organization. For plant species occupying discrete habitat niches, such as is the case with hairy prairie-clover, the relationship between pattern and process is important because plants are immobile, they exhibit a strong spatial structure through specific habitat preferences, and their dispersal is restricted (Freckleton and Watkinson, 2002).

Landscape pattern is generated by processes operating at various spatial and temporal scales (Urban et al., 1987). The relative influence of landscape pattern on a plant species will depend on the temporal and spatial scales at which particular processes of the plant community operate (Williams et al., 2006). As a result, plant populations can be categorized based on differences in spatial scale (Table 1.1). Freckleton and Watkinson (2002) came up with six categories to explain the differences in plant populations based on spatial scales of process and pattern. Eriksson (1996) similarly grouped plant populations into three discrete categories. In reality, a single species is capable of exhibiting characteristics of more than one population type (Eriksson, 1996).

Table 1.1: Discrete categorization of plant population types based on spatial scale of process and pattern as interpreted from Freckleton and Watkinson (2002), and Eriksson (1996).

Population Type	Description	Spatial Scale of Process	Spatial Scale of Pattern
Metapopulation	local populations exist as discrete patches which are spatially linked	Regional scale processes of colonization, migration, and extinction dominate	**landscape level analysis:** patterns of heterogeneity and homogeneity most important
Source-Sink	local populations exist as discrete patches which are isolated except from a few source patches	Regional scale processes of colonization, migration, and extinction dominate	**landscape level analysis:** patterns of heterogeneity and homogeneity most important
Regional Ensemble	local populations exist as discrete patches which are isolated from one another	Regional scale and local scale processes both important	**landscape level analysis:** patterns of homogeneity most important, heterogeneity secondary
Remnant Population	local populations exist as discrete patches which are isolated from one another; patches persist through a dormant seed bank	Regional scale and local scale processes both important	**landscape level analysis:** patterns of homogeneity most important, heterogeneity secondary
Spatially Extended Population	continuous populations exist within a large area of suitable habitat	Local scale processes dominate	**Patch level analysis:** patterns of homogeneity most important
Spatially Structured Local Population	a local population which is affected by spatial structure	Local scale processes dominate	**Patch level analysis:** patterns of heterogeneity and homogeneity most important
Spatially Unstructured Local Population	a local population that is continuous in space	Local scale processes dominate	**Patch level analysis:** patterns of homogeneity most important
Local Population	small-scale, i.e. a few m^2	Local scale processes dominate	**Patch level analysis:** patterns of homogeneity most important

Population processes can operate at both local and regional spatial scales. Local scale processes operate at the patch level while regional scale processes operate at the landscape level. Local scale processes are not in isolation from regional scale processes however, because regional availability of habitat, migration, and extinction will determine if a system of local populations can persist (Freckleton and Watkinson, 2002). At the patch level, localized processes of dispersal, establishment, and persistence can be represented as (Birth-Death) while regional scale processes of migration, colonization, and extinction can be represented as (Immigration-Emigration) at the landscape level (Figure 1.1). There is also an interaction of a population patch in landscape level processes which can be represented as ((Immigration+Emigration)-(Birth+Death)) (Freckleton and Watkinson, 2002; Kolb and Diekmann, 2005). This type of

population pattern represents a metapopulation model where there are N number of patches which are either occupied or unoccupied and the state of the whole metapopulation is a function of the states of localized populations (Heinz et al., 2006). Changes in the state of the metapopulation are a result of either changes in overall extinctions or changes in overall migration and colonization. Thus, the long term viability of a metapopulation depends on the number of individuals arriving at a patch and the number of individuals leaving a patch (Heinz et al., 2006).

While plants can exhibit many population patterns (Table 1.1), the metapopulation model is most appropriately applied to species living in highly fragmented landscapes (Eriksson, 1996; Hanski, 2004) where unsuitable habitat exceeds suitable habitat. This is characteristic of rare plants because they tend to have lower reproductive outputs, possess shorter lived seeds, are smaller in size, depend on seed output and seedling recruitment for population growth, and have a strong affiliation for a particular habitat niche (Landres et al., 1999; Kolb and Diekmann, 2005; Schleuning and Matthies, 2008; Vitt et al., 2009). The hairy prairie-clover population within the Dundurn-sand hills tends to exhibit a spatial pattern appearing like a metapopulation, forming systems of local population patches and non-occupied but potentially suitable habitat patches. Assuming a metapopulation pattern implies ongoing metapopulation processes, it can be important to consider landscape pattern at both the landscape level and patch level because metapopulation processes interact at these two scales.

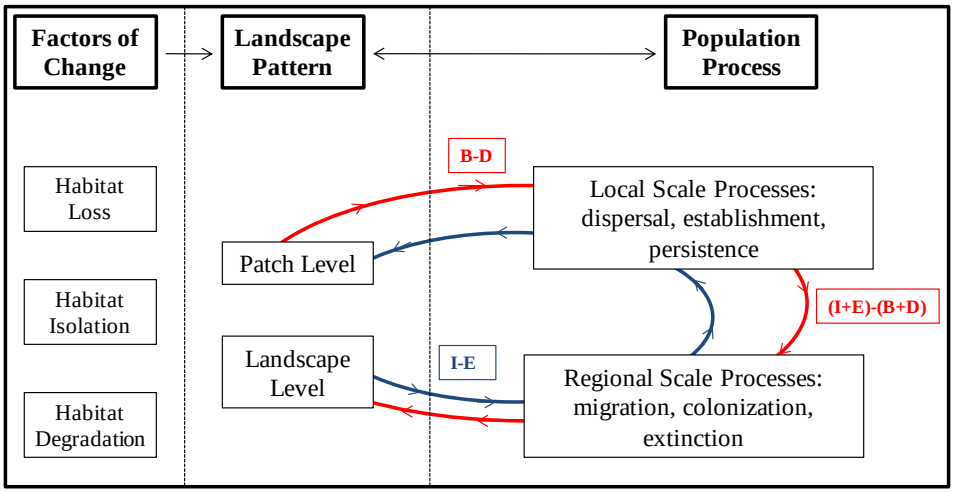


Figure 1.1: Relationship between pattern and process within a metapopulation model. Direct changes in landscape pattern at the patch level will affect the local plant population within that patch (red pathway). Local scale processes are indirectly affected as represented by the equation (B-D) where the ratio of births (dispersal and establishment: B) to deaths (persistence: D) will change. Between-patch interaction of local plant populations is represented by regional scale processes in the equation (I+E)-(B+D) where I represents immigration and E represents emigration. Direct changes in landscape pattern at the landscape level will affect the metapopulation within a region (blue pathway). Regional scale processes are indirectly affected as represented by the equation (I-E) where the ratio of immigration (colonization and extinction: I) to emigration (migration and extinction: E) will change.

1.3 Remote Sensing Applications to Ecology

With the growing popularity of remote sensing technology, it is now possible to fulfill many of the needs of environmental managers for conservation efforts (McDermid et al., 2005; Jobin et al., 2008). Current methods focus on capturing the homogeneity and heterogeneity within landscape pattern. Remote sensing technology focuses on grouping homogeneous subsets together to find discrete areas of commonality while landscape metrics analysis focuses on the spatial heterogeneity between these subsets (Gustafson, 1998; McDermid et al., 2005; Elmqvist et al., 2008).

Spatial habitat analysis builds a relationship between landscape pattern and population pattern to infer about ecological process. Landscape pattern observed on the ground can be represented in the spatial structure generated within a satellite image through land cover classification. Land cover classification is one of the most conventional outputs of remote sensing due to the ability of images to detect differences in reflectance for different vegetation and land use types. Land cover maps can be created based on homogeneous areas of vegetation cover. Consequently, habitat type maps can be derived from land cover maps based on the assumption that an area exhibiting similar vegetation cover is also likely to obtain homogeneous conditions for other environmental parameters (McDermid et al., 2005).

Landscape metrics analysis of a classified map can capture patterns of heterogeneity and homogeneity to show spatial distributions and associations of habitat types over time (Jobin et al., 2008). Metrics can be divided into six major categories which are able to represent ecological parameters of interest: patch area, patch shape, core area, isolation/connectivity, edge, and aggregation (Wang and Malanson, 2007). Specific patch metrics such as number of patches, patch area, edge-density, and mean shape index were found to be best at explaining habitat features for plant species (Uuemaa et al., 2009). Multiple regression analysis can be subsequently used to determine the metrics/landscape patterns that are significantly correlated with the distribution patterns of a plant population. Correlating landscape pattern with population pattern can allow for the inference about ecological processes that may be occurring within the landscape.

Landscape homogeneity and heterogeneity is a function of spatial configuration, temporal change, and both spatial and temporal scale. Landscape pattern and population pattern must be measured at the same spatial scale in order to infer about one another (Gustafson, 1998).

If landscape pattern and ecological process cannot be related within similar spatial scales, they cannot have an ecologically meaningful relationship. Analysis of remote sensing products may either fail to detect pattern in a landscape where pattern actually exists or may detect manufactured pattern in a landscape where no pattern exists (Li and Wu, 2004) because of the spatial scales at which data were acquired. Thus, when dealing with spatial analysis, identifying an appropriate scale of study is a vital precursor in order to produce ecologically meaningful remote sensing products (Lightowlers et al., 2008).

1.3.1 Using Remote Sensing to Solve Ecological Problems

Following the hierarchy theory presented by Noss (1990), it is suggested that aspects of landscape biodiversity exist at different organizational levels as represented in the composition, structure, and function of an ecosystem (Figure 1.2). Higher levels of organization can incorporate and constrain lower levels of organization such that environmental monitoring programs can select top levels of organization to represent the many aspects of biodiversity that fall below it (Urban et al., 1987; Noss, 1990). It has been well documented that remote sensing technology is capable of simulating the structural aspect of an ecosystem (McDermid et al., 2005). Using remote sensing to model the structural component of the landscape can provide a base for further research on the functional and compositional components of an area. Considering the hairy prairie-clover population within the study site exhibits a spatial pattern appearing like a metapopulation, the first two structural levels of organization should be considered because patterns at the patch level and landscape level can interact in a metapopulation. As well, studying landscape pattern and habitat pattern can provide a base for which to infer about lower levels of organization such as population pattern and genetic pattern (Noss, 1990). Therefore, the situation of this research within broader ecosystem studies is outlined in red (Figure 1.2).

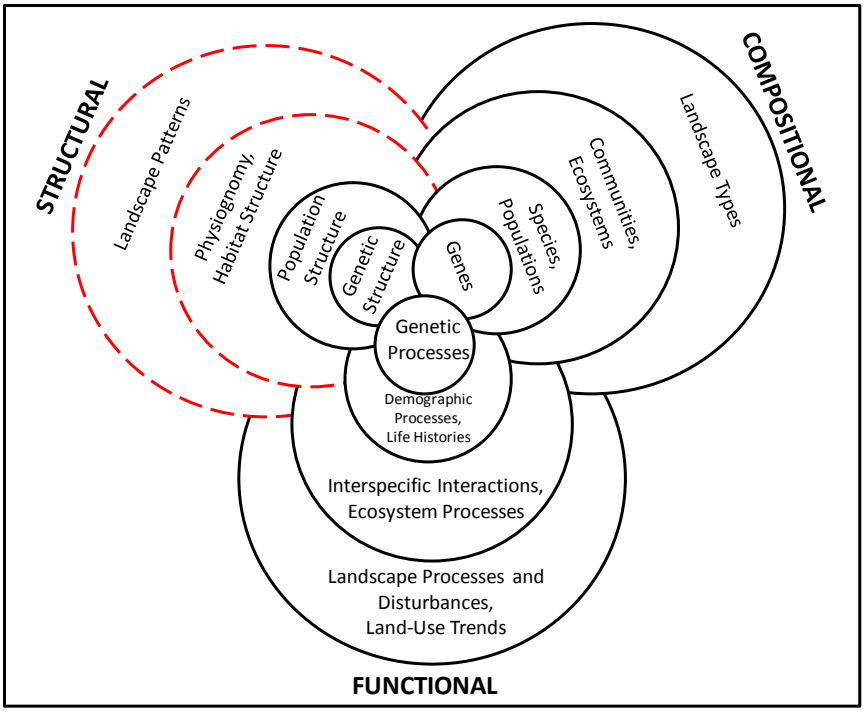


Figure 1.2: Hierarchy theory as adapted from Noss (1990). An ecosystem can be explained by its structure, function, and composition at hierarchical organizational levels. These organizational levels interact at varying scales such that higher organizational levels can constrain the behaviour of lower levels. This research will focus on structure through exploring landscape pattern and habitat pattern with the use of remote sensing technology. The situation of this research within the broader aspect of ecosystem studies is outlined in red.

1.4 Dalea villosa

Dalea villosa (hairy prairie-clover) is a member of the legume family. It is a perennial, reproducing mainly by seeds and flowering from late July to late August then setting seed in September. Within Canada, there are two closely related species within the same range as *Dalea villosa*: the white prairie-clover (*D. candidum*) and the purple prairie-clover (*D. purpereum*) (McGregor at el., 1986).

Hairy prairie-clover reaches the northern limits of its range in south central Saskatchewan and south western Manitoba, at approximately 51°N latitude (Smith, 1998). Its range extends southwards into the United States, to New Mexico, Texas, and Michigan. In Canada, the hairy prairie-clover was afforded protection under the Federal Government's Species at Risk Act in 2004, because of the limited area of occupancy, number of locations, and declining habitat quality. Within Canada, this species is found at two sites in Saskatchewan (Dundurn and Mortlach) and several sites in Manitoba (Treesbank, Camp Shilo, Spruce Woods Provincial Park, Boissevain, and the Lauder Sand hills) (Smith, 1998). In 2006, species numbers in the Dundurn sand hills area were estimated to be 109,556 plants (Godwin and Thorpe, 2007).

1.4.1 Habitat Conditions in Prairie Sand Dunes

In Canada, hairy prairie-clover is found within the Aspen Parkland and Mixed-grass Prairie of Saskatchewan and Manitoba. The dominant habitat requirement for this species appears to be an element of bare sand cover resulting in the plants confinement to geographical features such as parabolic sand dunes, stabilized blow-outs, dune depressions, and sand flats (Smith, 1998). Habitat preference is for south and west facing slopes with low angles, absence of an A-horizon, and low to partial vegetation cover with exposed bare sand (Godwin and Thorpe, 2007). This species' specific habitat requirement of sparsely vegetated sand means that it is also adapted to factors such as an annual water deficiency (limited moisture creating cohesion in soil), limited vegetation cover (limited root mass stabilizing soil), and sufficient wind speeds to entrain sand particles and facilitate aeolian erosion of dunes (Trenhaile, 2004).

Sand dunes of the Mixed-grass Prairie in Canada are comprised of aeolian deposits that were derived from glacio-fluvial sediments deposited in the region about 10,000 to 13,000 years before present (Hugenholtz and Wolfe, 2005). Several types of sand dune complexes are common in this region, namely active complexes, stabilized blow-outs, stabilized dunes (parabolic), and dune depressions (Hulett et al., 1966). Sand dune formation is sensitive to several factors including atmospheric conditions that limit sediment transport capacity (sand-moving power of the wind) and surface conditions that limit sediment supply and availability (vegetation cover) (Hugenholtz and Wolfe, 2005). For example, parabolic sand dunes occur when there is partial vegetation cover during dune formation limiting sediment availability, as was characteristic of dune formation within the southern Canadian prairies. Vegetation cover will stabilize part of the dune by decreasing the momentum of air flow, bounding soil with root systems, and trapping particles (Blanco et al., 2008). While the flanks of the dune become stabilized by vegetation, the slip face will remain relatively unvegetated (Blanco et al., 2008; Hugenholtz and Wolfe, 2005).

Wind is also one of the dominant factors shaping sand dune complexes. Winds in Saskatchewan are within the high energy category and wind speeds can be even greater than those found in desert basins. Generally, the highest wind speeds are experienced in April and May while wind speed reaches a minimum during July and August (Hugenholtz and Wolfe, 2005). For example, the formation of stabilized blow-outs is highly affected by wind as seen in the orientation of the saucer-shaped depressions to the direction of the prevailing wind. Wind

erosion of blow-outs will cease when the depression becomes deep enough to reduce the effective wind or when an unerodable surface is reached, allowing vegetation to take hold in the depression (Hulett et al., 1966). Thus, due to these and other environmental factors, sand dunes characteristically exhibit low soil fertility, low soil moisture holding capacity, and high erosion risk from wind (Wolfe, 1997). Hence, these partially stabilized sandy areas provide elements that are important to hairy prairie-clover habitat.

1.4.2 Threats to Dalea villosa Habitat

In Saskatchewan, only about 20% of the native Mixed-grass Prairie remains (Davis, 2004). There is evidence that since the mid 1900's overall dune activity has been decreasing within the Canadian prairies due to greater vegetation colonization compared to sand dune migration (Hugenholtz and Wolfe, 2005; Muhs and Wolfe, 1999). It has further been found that many sand dunes within southern Alberta, Saskatchewan, and Manitoba have vegetated at a rate of 10% to 20% per decade, nearly completely stabilizing over the last century (Vance and Wolfe, 1996). While it has been speculated that climate change is a dominant factor contributing to geomorphological changes in dune activity over the last century, other factors such as vegetation encroachment, invasive species encroachment, anthropogenic sources and land use changes, and loss of natural disturbance by bison and fire have been leading to the reduction in dune activity within the prairies (Wolfe et al., 1995; Vance and Wolfe, 1996; Smith, 1998; Hugenholtz and Wolfe, 2005). Similarly, it has also been identified that vegetation and invasive species encroachment of sand dunes, loss of grazing and fire disturbance pressures, and changes in land use patterns are the main threats to the survival of hairy prairie-clover (Smith, 1998).

1.5 Goal and Objectives

The goal of this thesis is to identify and spatially delineate areas of critical habitat for hairy prairie-clover within the range of a known population in the Dundurn-sand hills south of Saskatoon, Saskatchewan. The hairy prairie-clover population within the study site exhibits a key ecological aspect for consideration in critical habitat identification: habitat specialization for bare sand. Thus, it is hypothesised that the dominant factor limiting the abundance of hairy prairie-clover on the landscape is the spatial configuration of bare sand areas in relation to other landscape elements. Following Noss's (1990) hierarchical levels of organization it is assumed that landscape pattern and habitat pattern will exert the greatest influence on population pattern

(Figure 1.2). Understanding landscape pattern is important for understanding the effects of habitat change on rare flora populations such as hairy prairie-clover. Changes in landscape pattern can lead to changes in population pattern and thus changes in population processes which will threaten the long term survival of rare plant species that are habitat limited (Kolb and Diekmann, 2005; Kindlmann and Burel, 2008). Therefore, it is necessary to understand the spatial pattern of potential bare sand habitat for hairy prairie-clover within the study site landscape.

This research was divided into two specific objectives, first to look at landscape pattern and secondly to look at habitat pattern:

1) to investigate the spatial relationship between bare sand habitat for hairy prairie-clover and other land cover classes (Aspen, Shrub, Grasslands, Juniper)

2) to study the relationship between habitat configuration and hairy prairie-clover occurrence

1.5.1 Research Questions

To achieve the first objective, the desired output was a land cover classification of the study site. Remote sensing technology offers a more robust analysis of the spatial relationships between landscape elements than field based surveys because a much larger spatial and temporal scale can be used. However, the gap between using technology to generate information and the application of that information in an ecologically relevant context still remains largely un-traversed (McDermid et al., 2005). The scale dependency that exists between remote sensing technology and ecological relevance strongly affects the accuracy of a project. When dealing with spatial analysis, identifying an appropriate scale of study is a vital precursor however, this can be difficult to implement (Lightowlers et al., 2008). Therefore, objective one will specifically address two research questions:

1) What is the optimal spatial resolution for accurate and precise classification of bare sand areas?

2) What is the optimal temporal resolution for accurate land cover classification?

The desired output from objective two was a habitat suitability map for hairy prairie-clover. For species protected by federal laws, identification of critical habitat through the development of a habitat suitability map is necessary as a first step towards habitat conservation (Wolken et al., 2001). Creating a link between landscape pattern and population pattern can provide valuable insight into the long term viability of a population. Therefore, objective two focused on the relationship between the spatial pattern of bare sand areas and the occurrence of hairy prairie-clover. Objective two consisted of three specific research questions:

1) What is the spatial pattern of *intra-patch* characteristics for occupied and unoccupied bare sand habitat?

2) What is the spatial pattern of *inter-patch* characteristics for occupied and unoccupied bare sand habitat?

3) What spatial configuration of bare sand habitat best explains and predicts occupancy by hairy prairie-clover?

Acheiving these objectives will not only provide a basis for critical habitat identification of hairy prairie-clover, but will also provide a record of the spatial extent of bare sand and sand dunes within this region and time period. These records could be used in change detection analysis of sand dune stabilization, which has been hypothesised to be a valuable indicator of climate change and climate effects on the prairies (Hugenholtz and Wolfe, 2005; Wolfe, 1997; Scott and Suffling, 2000). As well, sand dune mapping can provide a record of habitat change over time for the hairy prairie-clover. Monitoring changes in habitat fragmentation can provide valuable insight into how a plant responds to changes in habitat, thus further aiding in land use management decisions.

1.6 Study Site

This study took place in south central Saskatchewan within the Dundurn and Rudy Rosedale Prairie Farm Rehabilitation Administration (PFRA) Community Pastures located at 51º 45’00”N and 106º 45’00”W (Figure 1.3). The area is a remnant patch of natural mixed-grass prairie and sand dune complexes bordered on the north by the Canadian Forces Base, 17-Wing Detachment

Dundurn, the east and south by mixed farm land, and the west by highway #219. The landscape is dominated by five main land cover types; aspen (*Populus tremuloides Michx.*), shrub (*Elaeagnus commutata Bernh. ex. Rydb*., *Prunus virginiana L. var. virginiana*, S*ymphoricarpos occidentalis Hook.*), juniper (*Juniperus horizontalis Moench.*), grassland, and sand dunes. This pasture is categorized as a low dunes ecosite with mainly stabilized sandy hills that have gentle to moderate slopes and areas between dunes that are relatively flat comprising a mixed-grass community (Thorpe, 2007).

The sand dunes found in the Dundurn area comprise four physiographic categories: active complexes, stabilized blow-outs, stabilized dunes, and dune depressions. Species most commonly associated with active complexes are *Psoralea lanceolota, Agropyron spp.,* and *Lygodesmia juncea*. Stabilized blow-outs are saucer-shaped depressions oriented in the direction of the prevailing wind and are most often vegetated by *Juniperus horizontalis, Carex heliophila,* and *Selaginella densa*. When dunes become stabilized, the resulting vegetation community is most dominated by *Stipa comata, Artemisia frigida,* and *Selaginella densa.* The difference between stabilized blow-outs and stabilized dunes is seen in the occurrence of low lying shrubs on the former which are less common on stabilized dunes. The dominant species in dune depressions, which occur on stabilized dunes, are *Stipa comata, Calamovilfa longifolia,* and *Agropyron spp.* (Hulett et al., 1966). Many population patches of hairy prairie-clover have been found in grassland and prostrate shrub stands within the study site, most commonly on low angled south and west facing slopes with partial vegetation cover and exposed sand (Godwin and Thorpe, 2007).

This PFRA pasture covers an area of 12 300 ha with a maximum elevation of 520 m above mean sea level. The Dundurn sand hills are characterized by annual temperature extremes with a mean annual temperature of 2.2ºC and a mean of 14.8ºC during the growing season from May to September (Environment Canada, 2008). The growing season is short with about 105 frost free days, high wind speeds (mean annual wind speed of 15.7 km hr^{-1}), high evaporation rates (mean annual evaporation rate of 650 mm), and low precipitation rates (mean annual precipitation rate of 350 mm) (Environment Canada, 2008). Prevailing winds in southern Saskatchewan are from the northwest resulting in a general dune orientation of northwest to southeast.

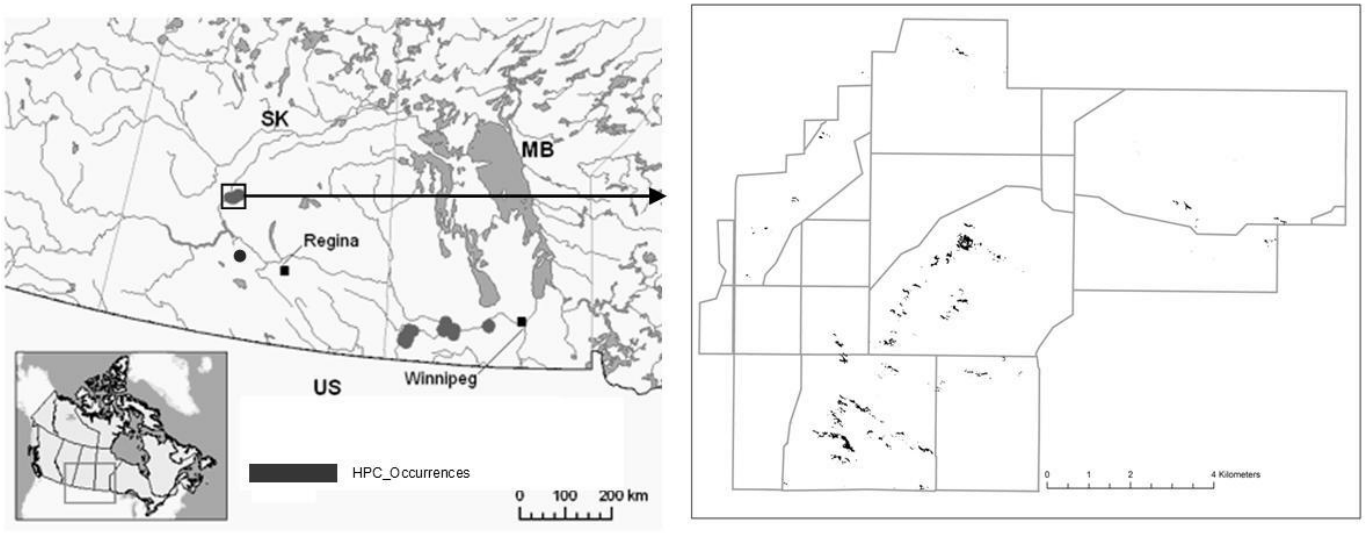


Figure 1.3: Dundurn and Rudy Rosedale Prairie Farm Rehabilitation Administration Community Pastures, Saskatchewan. The Canadian distribution of hairy prairie-clover is shown on the left (Canadian Wildlife Service, 2004). A close-up of the study site is shown on the right with pasture boundaries outlined in grey. The known distribution of the population is shown by the black polygons.

1.7 Organization of Thesis

This thesis has been divided into four chapters. The first chapter consists of a literature review in order to situate this research within the context of broader ecosystem studies. The second chapter focuses on the first objective of this research, mainly investigating landscape pattern through land cover classification of the study site. The third chapter deals with the second objective of this research, focusing on habitat pattern and habitat suitability mapping for hairy prairie-clover. The final chapter of this thesis brings together the findings of chapters two and three, and discusses the significance, limitations, and future research opportunities of this thesis.

1.8 References

Alados, C., Navarro, T., Komac, B., Pascual, V., Martinez, F., Cabezudo, B., Pueyo, Y., 2009. Do vegetation patch spatial patterns disrupt the spatial organization of plant species? *Ecological Complexity* 6: 197-207.

Blanco, P., Rostagno, C., Valle, H., Beeskow, A., Wiegand, T., 2008. Grazing impacts in vegetated dune fields: predictions from spatial pattern analysis. *Rangeland Ecol Manage* 61: 194-203.

Canadian Wildlife Service, 2004. Species at Risk Public Registery: Canadian distribution of the hairy prairie-clover. URL: http://www.sararegistry.gc.ca/species/speciesDetails_e.cfm?sid=533 [Last accessed: October 29, 2010].

Canadian Wildlife Service, 2005. Species at Risk Act Implementation Guidance: Draft technical guidelines for identifying critical habitat. *Environment Canada, Ottawa, Ont.*

Davis, S., 2004. Area sensitivity in grassland passerines: Effects of patch size, patch shape, and vegetation structure on bird abundance and occurrence in southern Saskatchewan. *The Auk* 121: 1130-1145.

Elmqvist, B., Ardo, J., Olsson, L., 2008. Land use studies in drylands: an evaluation of object oriented classification of very high resolution panchromatic imagery. *International Journal of Remote Sensing* 29: 7129-7140.

Environment Canada, 2008. National climate data and information archive: Canadian climate normals 1971-2000. URL: http://_www.climate.weatheroffice.ec.gc.ca/climate_normals [Last accessed: May 13, 2009].

Eriksson, O., 1996. Regional dynamics of plants: a review of evidence for remnant, source-sink and metapopulations. *Oikos* 77: 248-258.

Fahrig, L. and Merriam, G., 1994. Conservation of fragmented populations. *Conservation Biology* 8: 50-59.

Fahrig, L., 2001. How much habitat is enough? *Biological Conservation* 100: 65-74.

Franks, A., Yates, C., Hobbs, R., 2009. Defining plant functional groups to guide rare plant management. *Plant Ecology* 204: 207-216.

Freckleton, R. and Watkinson, A., 2002. Large-scale spatial dynamics of plants: metapopulations, regional ensembles and patchy populations. *Journal of Ecology* 90: 419-434.

Godwin, B. and Thorpe, J., 2007. Targeted surveys for plant species at risk in Elbow, Dundurn and Rudy-Rosedale AAFC-PFRA Pastures, 2006. Agriculture and Agri-Food Canada-Prairie Farm Rehabilitation Administration. *Saskatchewan Research Council Pub. No. 11997-1E07.*

Gustafson, E., 1998. Quantifying landscape spatial pattern: What is the state of the art? *Ecosystems* 1: 143-156.

Hanski, I., 2004. Metapopulation theory, its use and misuse. *Basic and Applied Ecology* 5: 225-229.

Heinz, S., Wissel, C., Frank, K., 2006. The viability of metapopulations: individual dispersal behaviour matters. *Landscape Ecology* 21: 77-89.

Hodges, K. and Elder, J., 2008. Critical habitat designation under the US endangered species act: How are biological criteria used? *Biological Conservation* 141: 2662-2668.

Hoekstra, J., Fagan, W., Bradley, J., 2002. A critical role for critical habitat in the recovery planning process? Not yet. *Ecological Applications* 12: 701-707.

Hugenholtz, C. and Wolfe, S., 2005. Recent stabilization of active sand dunes on the Canadian

prairies and relation to recent climate variations. *Geomorphology* 68: 131-147.
Hulett, G., Coupland, R., Dix, R., 1966. The vegetation of dune sand areas within the grassland region of Saskatchewan. *Canadian Journal of Botany* 44: 1307-1331.
Jobin, B., Labrecque, S., Grenier, M., Falardeau, G., 2008. Object-based classification as an alternative approach to the traditional pixel-based classification to identify potential habitat of the Grasshopper Sparrow. *Environmental Management* 41: 20-31.
Kindlmann, P. and Burel, F., 2008. Connectivity measures: a review. *Landscape Ecology* 23: 879-890.
Kolb, A. and Diekmann, M., 2005. Effects of life-history traits on responses of plant species to forest fragmentation. *Conservation Biology* 19: 929-938.
Landres, P., Morgan, P., Swanson, F., 1999. Overview of the use of natural variability concepts in managing ecological systems. *Ecological Applications* 9: 1179-1188.
Li, H. and Wu, J., 2004. Use and misuse of landscape indices. *Landscape Ecology* 19: 389-399.
Lightowlers, C., Nelson, T., Setton, E., Keller, C., 2008. Determining the spatial scale for analysing mobile measurements of air pollution. *Atmospheric Environment* 42: 5933-5937.
McDermid, G., Franklin, S., LeDrew, E., 2005. Remote sensing for large area habitat mapping. *Progress in Physical Geography* 29: 449-474.
McGregor, R., Barkley, T., Great Plains Flora Association (U.S.), 1986. *Flora of the Great Plains.* Kansas: University Press of Kansas.
Muhs, D. and Wolfe, S., 1999. Sand dunes of the northern great plains of Canada and the United States. In: Holocene Climate and Environmental Change in the Palliser Triangle. D.S. Lemmen and R.E. Vance (Eds.), Geological Survey of Canada, Bulletin 534: 183-197.
Naujokaitis-Lewis, I., Curtis, J., Arcese, P., Rosenfeld, J., 2009. Sensitivity analyses of spatial population viability analysis models for species at risk and habitat conservation planning. *Conservation Biology* 23: 225-229.
Noss, R., 1990. Indicators for monitoring biodiversity: A hierarchical approach. *Conservation Biology* 4: 355-364.
Schleuning, M. and Matthies, D., 2008. Habitat change and plant demography: Assessing the extinction risk of a formerly common grassland perennial. *Conservation Biology* 23: 174-183.
Scott, D. and Suffling, R., 2000. Climate change and Canada's national park system: a screening level assessment. *Environment Canada,* En56-155/2000E, 183 pp.
Smith, B., 1998. COSEWIC status report on the hairy prairie-clover in Canada. *Environment Canada.*
Thorpe, J., 2007. Saskatchewan Rangeland Ecosystems, Publication 1: Ecoregions and Ecosites. Saskatchewan prairie conservation action plan. *Saskatchewan Research Council Pub. No. 11881-1E07.*
Trenhaile, A., 2004. *Geomorphology: a Canadian perspective 2nd edition*. Oxford University Press, Canada.
Urban, D., O'Neill, R., Shugart, H., 1987. A hierarchical perspective can help scientists understand spatial patterns. *BioScience* 37: 119-128.
Uuemaa, E., Antrop, M., Roosaare, J., Marja, R., Mander, U., 2009. Landscape metrics and indices: An overview of their use in landscape research. *Living Reviews in Landscape Research* 3: 5-28.
Vance, R. and Wolfe S., 1996. Geological indicators of water resources in semi-arid environments: southwestern interior of Canada. In: Geoindicators: Assessing rapid environmental changes in earth systems. Berger, A.R. and Iams, W.J. (Eds.).

A.A. Balkema, p. 251-263.
Vitt, P., Havens, K., Kendall, B., Knight, T., 2009. Effects of community-level grassland management on the non-target rare annual *Agalinis auriculata*. *Biological Conservation* 142: 798-805.
Wang, Q. and Malanson, G., 2007. Patterns of correlation among landscape metrics. *Physical Geography* 28: 170-182.
Williams, N., Morgan, J., McCarthy, M., McDonnell, M., 2006. Local extinction of grassland plants: The landscape matrix is more important than patch attributes. *Ecology* 87: 3000-3006.
Wimberly, M., 2006. Species dynamics in disturbed landscapes: When does a shifting habitat mosaic enhance connectivity? *Landscape Ecology* 21: 35-46.
Wolfe, S., 1997. Impact of increased aridity on sand dune activity in the Canadian prairies. *Journal of Arid Environments* 36: 421-432.
Wolfe, S., Huntley, D., Ollerhead, J., 1995. Recent and late Holocene sand dune activity in southwestern Saskatchewan. *Geological Survey of Canada* 131-140.
Wolken, P., Sieg, C., Williams, S., 2001. Quantifying suitable habitat of the threatened western prairie fringed orchid. *Journal of Range Management* 54: 611-616.

2. Optimal Spatial and Temporal Resolution for Object-Oriented Classification of Prairie Sand Dune Landscapes

2.1 Abstract

Land cover classification is a key first step in habitat identification for hairy prairie-clover due to the plants specialization for bare sand environments, making bare sand identification and delineation of particular importance. However, when dealing with spatial analysis, identifying an appropriate scale of study is a vital precursor due to the scale dependency that exists between remote sensing products, such as land cover maps, and ecological relevance. The objective of this research was to investigate the spatial relationship between bare sand habitat for hairy prairie-clover and other land cover classes (aspen, shrub, grassland, juniper). To achieve this objective, a land cover classification of the study site at an appropriate spatial and temporal resolution is required. Spectral data was collected in the field for each land cover class in the early, middle, and late growing season of 2009 to determine the optimal temporal and spatial resolution for the remote sensing data set. Results showed that in the early and late growing season, the most spectral separability between classes occurred. The red, NIR, MIR, and panchromatic bands were most useful for differentiating between land cover classes. Wavelet analysis of the field data revealed an optimal spatial resolution between 2-5 m. Therefore, the remote sensing data set consisted of a SPOT5 panchromatic image at 2.5 m spatial resolution from April and 3 SPOT5 multispectral images at 10 m spatial resolution from May, July, and August. Object-oriented classification methods, particularly the segmentation procedure, can achieve improvements in patch delineation over pixel-based methods because it focuses on the geometric quality of objects. Object-oriented classification of the study site resulted in an overall accuracy of 79% with a user's and producer's accuracy of 85% for bare sand patches. Grassland comprised the matrix of the area covering 45.5% of the study site. Aspen and shrub were the most dominating landscape elements comprising 25.5% and 19.2% of the study site respectively. Bare sand made up only 6.0% of the study site while juniper was the least persistent class comprising only 2.7% of the study site.

2.2 Introduction and Background

Remote sensing and land cover classification of prairie landscapes is important for extracting such features as above ground biomass for livestock forage availability, leaf area index, or the spatial distribution of rare habitat types such as bare sand, on a large scale (Zhang and Guo, 2007). However, the gap between using technology to generate information and the application of that information in an ecologically relevant context still remains largely un-traversed. This can be described as the "technology-ecology interface" where gaps in understanding exist between the tools/technology and the ecological relevance of the information generated (McDermid et al., 2005). Scale is arguably one of the most important parameters in remote sensing because its appropriate application has the ability to traverse the technology-ecology gap. The scale dependency that exists between remote sensing products, such as land cover maps, and ecological relevance strongly affects the accuracy of a project, therefore identifying an appropriate scale of study is a vital precursor (Lightowlers et al., 2008). Despite this, most studies still fail to integrate the use of appropriate spatial and temporal resolutions into their data sets (Gustafson, 1998). As a result, remote sensing products may either fail to detect pattern in a landscape where pattern actually exists or may detect manufactured pattern in a landscape where no pattern exists (Li and Wu, 2004) because of the scales at which data were acquired.

2.2.1 Spatial Resolution

The precision with which land cover classes are spatially delineated and the accuracy in their spatial location is important when mapping prairie landscapes. For example, spatial records of bare sand and sand dune extent can be important in change detection analysis of sand dune stabilization which has been hypothesised to be a valuable indicator of climate change effects on the prairies (Hugenholtz and Wolfe, 2005; Wolfe, 1997; Scott and Suffling, 2000). As well, mapping bare sand patches can provide a record of habitat change over time for many rare prairie plants such as hairy prairie-clover. Accuracy in spatial delineation will depend on image spatial resolution.

Relationships between spatial resolution and spatial structure are central to remote sensing land cover classification (Ju et al., 2005). Spatial structure refers to landscape patterns observed within an image as a result of landscape homogeneity and heterogeneity. The ability to detect landscape homogeneity and heterogeneity depends on the size of image pixels (Zhang and Guo, 2007). Therefore, in order to accurately represent landscape pattern observed on the ground

in the spatial structure generated from an image, the spatial resolution of the image should be directly related to the spatial scale that shows the dominant variations within landscape pattern. Zhang and Guo (2007) identify two methods for relating spatial scales to image spatial resolution: wavelet transform and semivariogram. The use of wavelet transforms for scale analysis has become increasingly popular over the semivariogram because it can produce hierarchical frequency components at multiple spatial resolutions (Chen and Blong, 2003). It has been adequately documented that the spatial variation observed in landscapes is a function of the variability in the scale at which abiotic and biotic factors operate (He et al., 2007). Therefore, a landscape can have more than one significant spatial scale. For example, Rahman et al. (2003) found that patterns such as leaf area index varied at scales of 12 m while topography varied at scales around 36-40 m in the southern California grasslands. Therefore, understanding the relationship between the spatial structure present in an image and the image's spatial resolution is imperative to producing an ecologically relevant product (Chen and Blong, 2003; Zhang and Guo, 2007).

2.2.2 Temporal Resolution

Several problems inherent in grassland classification still exist: compared to trees or shrubs, grass species are much smaller than any spatial resolution currently available; unlike crops, spatial variation in coverage will occur within a certain grassland type; and spatial variation in the types of species present within a certain grassland type will occur (Herrera et al., 2009). Grassland classification is also complicated by the existence of standing dead materials, litter, biological crust, and bare soil affecting the spectral signature of land cover classes (Zhang and Guo, 2007). Therefore, it is rare to find pixels that are "pure" in grassland classifications (Herrera et al., 2009), making spectral separability between classes an issue.

The reflectance of vegetation is variable over time due to leaf production and senescence occurring at different times throughout the growing season for different vegetation types. Therefore, image acquisition should occur at several times during the growing season when contrasts in vegetation are exaggerated enough to be captured by differences in spectral reflectance (Lucas et al., 2007). This will enhance spectral variability between land cover categories and increase classification accuracy. For example, Jobin et al. (2008) used imagery acquired from spring and summer to better discriminate between annual and perennial crops when classifying potential habitat for the Grasshopper Sparrow in SW Quebec, Canada. Smith

and Fuller (2001) also found that with just two images, one from the early growing season and one from the late growing season, the ability to distinguish between land cover types was greatly enhanced. Therefore, optimizing the temporal resolution of the data set can increase spectral separability between classes to address some of the issues of grassland classification.

2.2.3 Classification Methods

Traditional methods of pixel-based land cover classification focus on analysis of spectral information from single pixels which may consequently subdivide the landscape into an arbitrary grid system not representative of the actual spatial structure present (Smith and Fuller, 2001). Comparatively, object-oriented classification uses the segmentation procedure to first create objects that are homogeneous in both spectral and spatial aspects, and then classifies the objects (Platt and Rapoza, 2008). The segmentation algorithm is not solely based on the spectral values of a single pixel but, also includes aspects such as texture and the pixels spatial relationship with surrounding pixels (Jobin et al., 2008). The focus of the segmentation procedure on boundary delineation and the geometric quality of a class (Moller et al., 2007) can help improve the accuracy with which land cover classes are spatially delineated from traditional pixel-based methods which do not typically focus on boundary delineation.

The quality of a segmentation depends on image spatial resolution and user-defined parameter settings both of which should be scaled to capture the spatial variation observed within the landscape on the ground (Moller et al., 2007; Platt and Rapoza, 2008). The scale parameter in image segmentation controls the maximum allowable amount of heterogeneity within an object by changing the size of the object in relation to image spatial resolution (Platt and Rapoza, 2008). Spectral and shape heterogeneity within an object can be determined by weighting the color and shape parameters between zero and one. Spectral heterogeneity is the sum of standard deviations of each image band while shape heterogeneity is defined by compactness and smoothness (Platt and Rapoza, 2008). Smoothness is the ratio of an object's perimeter to the perimeter of the object's bounding box. Compactness is the ratio of an object's perimeter to the square root of the number of pixels in the object (Elmqvist et al., 2008). The ability of the user to choose the maximum amount of spectral heterogeneity within an object can help improve spectral separability between classes to address some issues inherent in grassland classification.

2.2.4 Objective

The objective of this research is to investigate the spatial relationship between bare sand and other land cover classes (Aspen, Shrub, Grasslands, Juniper (Ch.1)) in a complex landscape. The purpose of this research is to analyse landscape pattern to gain an understanding of the spatial structure of prairie sand dune complexes and to evaluate bare sand habitat in context to surrounding landscape elements (Figure 1.2). To achieve the first objective, a land cover classification of the study site at an appropriate spatial and temporal resolution is required.

2.3 Methods

A land cover classification of the study site using SPOT5 panchromatic and SPOT5 multispectral imagery in conjunction with object-oriented classification methods was undertaken (Figure 2.1). Analysis of class scaled metrics of the different land cover types in the classification can answer objective one.

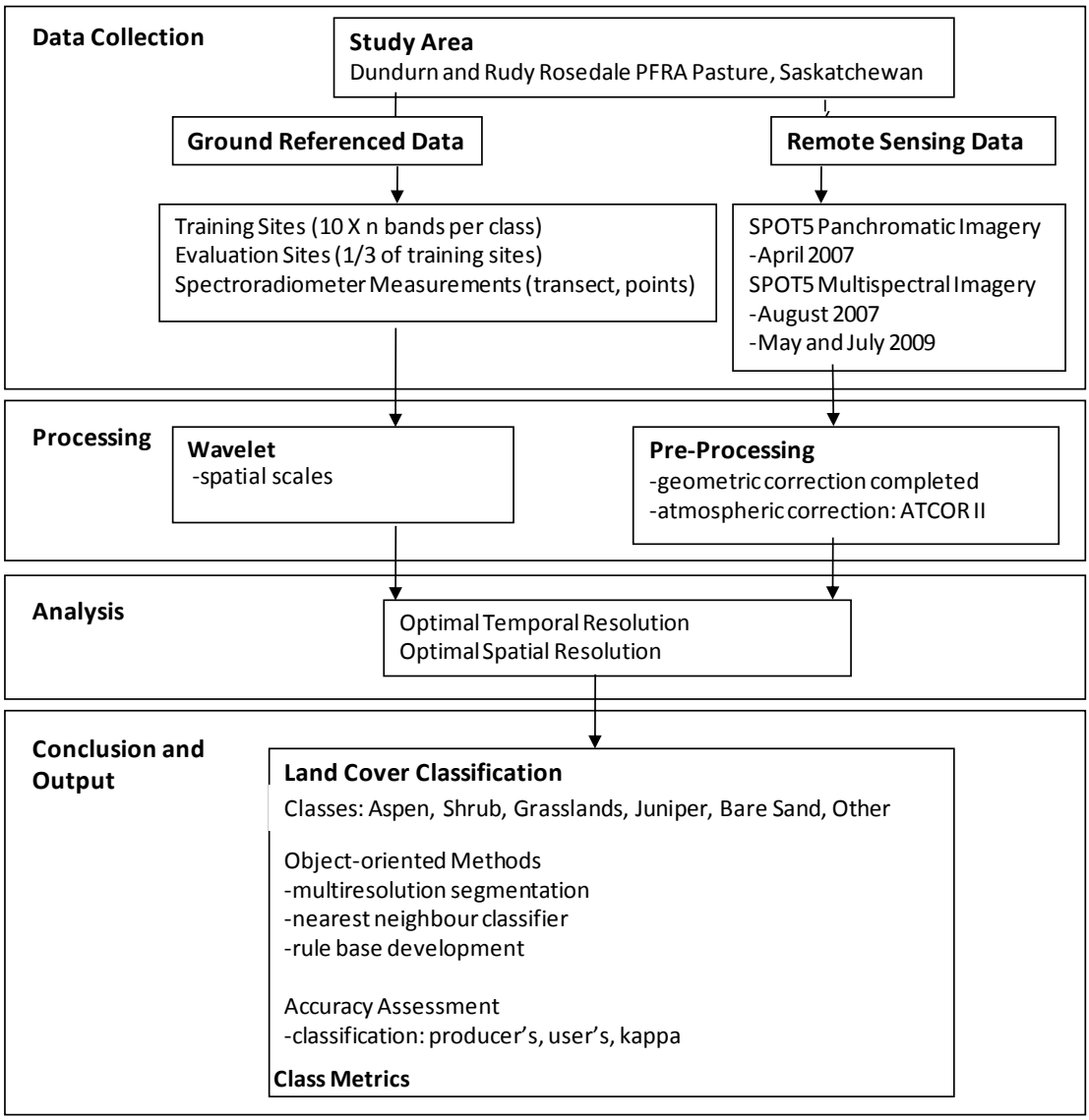


Figure 2.1: Flow chart of methods for objective one.

2.3.1 Remote Sensing Data Set

Four SPOT5 images were selected for the land cover classification of the study site. A SPOT5 panchromatic image (wavelengths 490-690 nm) at 2.5 m spatial resolution was acquired on April 13, 2007. This imagery was selected for its high spatial resolution. Three SPOT5 multispectral images (green, red, NIR, and MIR bands) at 10 m spatial resolution were acquired on May 27, 2009; July 12, 2009; and August 16, 2007. These images were selected to gain the needed spectral bands and temporal resolution for the data set. The imagery was obtained orthorectified and georeferenced to the NAD 1983 UTM Zone 13N projection. When using imagery acquired from different dates, atmospheric correction is necessary due to differences from atmospheric

effects such as clouds or haze, and illumination effects such as changes in sun angle (Song et al., 2001, McDermid et al., 2005). Atmospheric correction was completed using ATCORII within PCI Geomatica software.

2.3.2 Ground Referenced Data Set

Land cover classification requires the collection of ground truthed data: both training sites to train the classification algorithm to recognize specific land cover types and validation sites to assess the accuracy of the classification. Conventionally, the number of training sites needed is ten times the number of bands used, for each class (cited in Jensen, 2005). Four bands were used in the classification (Red, NIR, MIR, Panchromatic) thus, fourty training sites for each class were collected in the field. One third of the training sites (fourteen per class) were separated for use as validation sites post classification (cited in Jensen, 2005). Training sites were randomly selected and chosen to cover a range of conditions that would appear in each class. For example, the bare sand class consisted of sites that were pure sand to partially vegetated sand. GPS waypoints were taken in the center of patches to mark the geographical locations of the different land cover classes (Figure 2.2).

Spatial and temporal resolution analyses also require the collection of detailed ground-based measurements. In-situ spectroradiometer measurements along with biophysical parameters were collected for classes of bare sand , juniper, grassland, and shrub. Class aspen was excluded because it is not possible to collect above canopy spectral reflectance with a hand held spectroradiometer. One, 508 m long transect oriented to bisect sand dune complexes was set up in the field (Figure 2.2). Measurements were taken along the transect at 4 m intervals (2^7 sampling points). Spectroradiometer measurements were collected for wavelengths 350-2500 nm. Measurements were taken at 1 m above the ground, between 10:00 am and 2:00 pm to avoid distortions from sun angle, and on clear days to avoid shading by clouds.

Dominant species present and estimated percent vegetation cover were based on quadrats placed at 4 m intervals along the transect. For cover types bare sand and grassland, quadrats were 50 cm by 50 cm. Four quadrat measurements were taken at each point and then averaged in order to capture the spatial variation in grassland and bare sand cover. For cover types juniper and shrub, quadrats were 2 m by 2 m and only one quadrat measurement was taken at each point. Measurements were collected in the early, middle, and late growing seasons within 2009 in an attempt to capture temporal variation in spectral response curves of land cover types.

To supplement transect data for spatial and temporal resolution analyses, twenty additional sample points for each class were randomly chosen within the study site to repeat the above spectroradiometer and vegetation cover measurements. Equal stratified sampling (Hirzel and Guisan, 2002) was used to randomly select an identical number of points within each land cover class. Points were randomly generated within ArcMap9.2 using Hawth's Tools method: random selection within subsets. Fifty points within each land cover class were generated with the tool. From this, twenty points for each class were selected based on logistics: travel efficiency, 1 km proximity to roads, and ease of access. Please refer to Figure 2.2 for field design and study site selection.

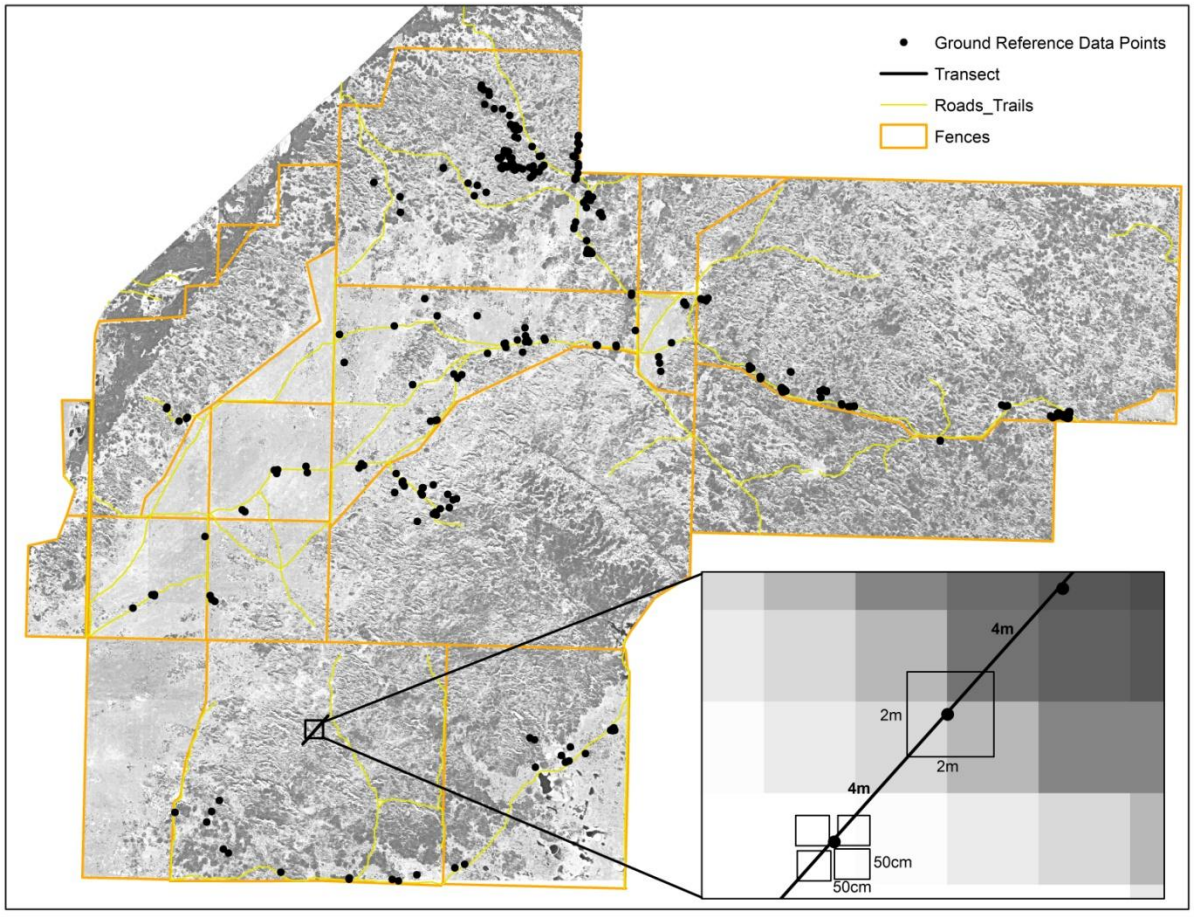


Figure 2.2: Field design and study site selection. The study site is represented by SPOT5 panchromatic imagery in the early growing season. The transect is represented by the black line and ground referenced data points are represented by the black points.

2.3.3 Temporal Resolution Analysis

Spectral response curves were constructed from field collected spectroradiometer measurements for each land cover class in the early, middle, and late growing seasons of 2009. Corresponding to the band width in the SPOT5 multispectral imagery, wavelengths 400-499 nm, 500-590 nm, 610-680 nm, 790-890 nm, and 1580-1750 nm were averaged to represent the blue, green, red, near infra-red (NIR), and mid infra-red (MIR) bands respectively. Each land cover class was represented as an average of the twenty random points for that class.

The coefficient of variance (CV) can be used to determine the relative spectral homogeneity within classes by determining dispersion of the data (Zhang and Guo, 2007; Zar, 1999). Maximizing within class homogeneity will better highlight between class variability allowing for better spectral separability between classes (Zhang and Guo, 2007). The CV was calculated as in Equation 2.1 across the twenty study sites for each class. Each class was calculated separately for the blue, green ,red, NIR, and MIR bands in the early, middle, and late growing seasons of 2009. The CV can be written as

$$CV = \frac{s}{\bar{x}} \tag{2.1}$$

where *s* is the standard deviation and ẋ is the mean of the twenty study sites for each class.

2.3.4 Spatial Resolution Analysis

Wavelet analysis can capture the spatial variability that occurs across a landscape at multiple scales by decomposing the signal (spectral reflectivity) into a series of frequency components at successively coarser spatial scales (Chen and Blong, 2003; He et al., 2007). Sampling of the signal should occur as fractional powers of two to limit edge effects and speed up the Fourier transform as data is assumed to be cyclical in a wavelet (Torrence and Compo, 1998 ;Chen and Blong, 2003). The result of a wavelet analysis is also highly dependent on the type of wavelet used and one should choose the wavelet function that best represents the type of data present in the signal (Torrence and Compo, 1998; Labat, 2005). To determine the significant spatial scales in spectral reflectivity on the ground, spectroradiometer measurements from the field transect with 2^7 sampling points spaced 4 m apart were used. Measurements collected in the late growing season (August 2009) were analyzed because the most spectral homogeneity was observed within classes at this time (Table 2.3). Minimizing within class heterogeneity will better highlight between class variability for use in wavelet analysis (Zhang and Guo, 2007).

Wavelengths 490-690 nm (SPOT5 panchromatic band) and 610-680 nm (SPOT5 red band) were analysed because high reflectance in these bands are characteristic of bare sand (Figure 2.11). The morlet wavelet mother function presented in Equation 2.2 was used to analyse the transect data because it is complex and better captures oscillatory behaviour (Torrence and Compo, 1998). The morlet wavelet can be written as

$$\psi a, b(t) = \frac{1}{\sqrt{a}}\psi(\frac{t-b}{a}) \quad (2.2)$$

where *a* is a contraction or dilation of the wavelet function corresponding to different scales, and *b* is the spatial shift of the wavelet function (He et al., 2007). The wavelet variance was analyzed to determine its statistical significance against a red noise data series at the 95% confidence interval (He et al., 2007). The free Matlab-software provided by C. Torrence and G. Compo (1998) available at URL: http://paos.colorado.edu/research/wavelets/ was used.

According to the sampling theorem presented by Mcgrew and Monroe (2000), in order to spatially capture an object one must sample at least at half the width of the object. Therefore, the optimum spatial resolution (size of a pixel/sampling object) should be at least one half of the semivariogram lag distance; the distance above which pixels/sampling objects are not related (Rahman et al., 2003). Since semivariograms are calculated to half of the maximum distance between points (half of a wavelet oscillation), sampling should therefore be one half of the lag distance of the semivariogram or one quarter of the wavelet scale (Guo et al., 2001; Rahman et al., 2003). To convert significant spatial scales to spatial resolution, one quarter of the wavelet scale was taken (Guo et al., 2001; Rahman et al., 2003; He et al., 2007).

2.3.5 Image Segmentation

The quality of a segmentation depends on image spatial resolution and user-defined parameter settings (Moller et al., 2007; Platt and Rapoza, 2008). The same transect was used to extract reflectance values from the SPOT5 panchromatic image. The semivariogram was computed to determine the lag distance at which pixels in the image were no longer spatially autocorrelated (Lightowlers et al., 2008). Equation 2.3 shows that semivariance is the averaged sum of squared differences in attribute values for data pairs that are a defined distance apart (Lightowlers et al., 2008). It can be written as

$$y(d) = \frac{1}{2n(d)}\sum_{s_i - s_j}(Z_i - Z_j)^2 \quad (2.3)$$

where Z_i and Z_j are the attribute values of data pairs S_i and S_j, *d* is the distance between data pairs, and *n* is the number of data pairs (Lightowlers et al., 2008). Reflectance is a spatially continuous variable given the spatial resolution of the imagery such that the spatial variation between reflectance values of any two pixels will depend on the lag distance (Rahman et al., 2003). The size of a sampling object should be one half of the semivariogram lag distance; the distance above which sampling objects are not related (Rahman et al., 2003). Therefore, the average distance between neighbouring image object centroids should be near to the lag distance of the semivariogram so as to group spatially autocorrelated pixels together (Karl and Maurer, 2010). Adjusting the scale parameter during image segmentation will adjust the size of image objects and the distance between object centroids.

To test the effect of scale in image segmentation, the shape and color parameters within the segmentation algorithm were held constant at 0.1 and 0.9 respectively. In natural landscapes where ecotones occur, object shapes can be difficult and sometimes ecologically irrelevant to define. In such cases, it may be more beneficial to capture the spatial variation in spectral reflectivity across a landscape because spectral heterogeneity can better represent the gradient between land cover types. Segmenting image objects based on spectral heterogeneity will allow for better correlation between field variables and remote sensing products which rely on the relationships inherent in spectral signatures, for example leaf area index (LAI) and normalized difference vegetation index (NDVI) (He et al., 2007). Therefore, the weight of the decision was put on the color parameter. The smoothness and compactness parameters were held constant at 0.5. Scale parameters of five, ten, fifteen, and twenty were tested for segmentation of the SPOT5 panchromatic image. Ten locations were randomly selected within the imagery to identify ten test objects. The average Euclidean distance between neighbouring image object centroids was calculated across the ten test objects. Iterations were run for each scale parameter. The optimum scale parameter was chosen when the average distance between neighbouring image object centroids was near to the lag distance of the semivariogram.

2.3.6 Land Cover Classification

A multi-temporal, multi-resolution land cover classification was carried out using object-oriented methods. Class boundary delineation was completed using a multiresolution segmentation algorithm at a scale parameter of ten, with shape and compactness weighted at 0.1 and 0.5 respectively. The weight of the decision was put on the spectral heterogeneity captured within an

object because spectral heterogeneity can better represent the gradient between land cover types than concrete shapes. Therefore, color was weighted at 0.9 during segmentation (shape at 0.1). The panchromatic band was given the most weight during segmentation to take advantage of the high spatial resolution. The NIR bands from the July 2009 and August 2007 multispectral images were also included in the segmentation to take advantage of the spectral separability between classes.

Following segmentation, the landscape was classified in to the five main land cover types present in the study site: aspen, shrub, juniper, grassland, and bare sand with a sixth class, other, added during post classification editing. A nearest neighbour classifier was set up using the twenty six training data points collected in the field for each class and class membership values were assigned based on the mean reflectance within the panchromatic, red, NIR, and MIR bands of an object. A hierarchical rule set was developed to maximize spectral and shape separability between classes (Figure 2.3). For each image, reflectance values were extracted for the training data points and the average reflectance was calculated for each class. From this, spectral curves of the land cover classes were constructed for each image (Figure 2.4). Threshold values, for which to base the rule sets from, were derived from these.

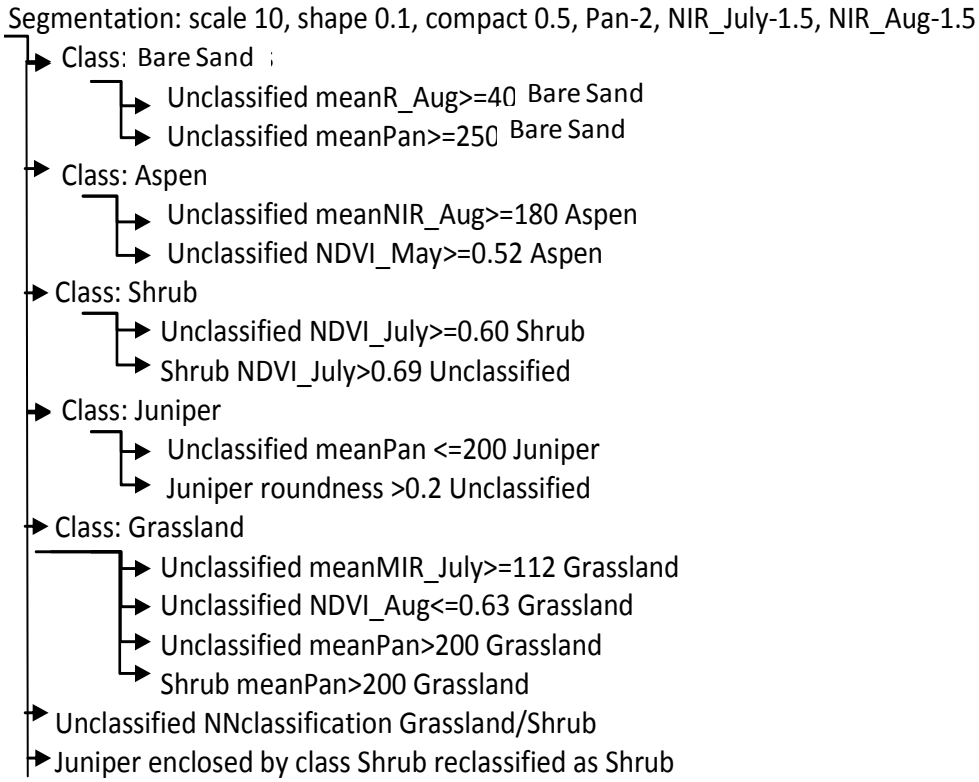


Figure 2.3: Rule set for object-oriented land cover classification using SPOT5 panchromatic imagery from April and three SPOT5 multispectral images from May, July, and August.

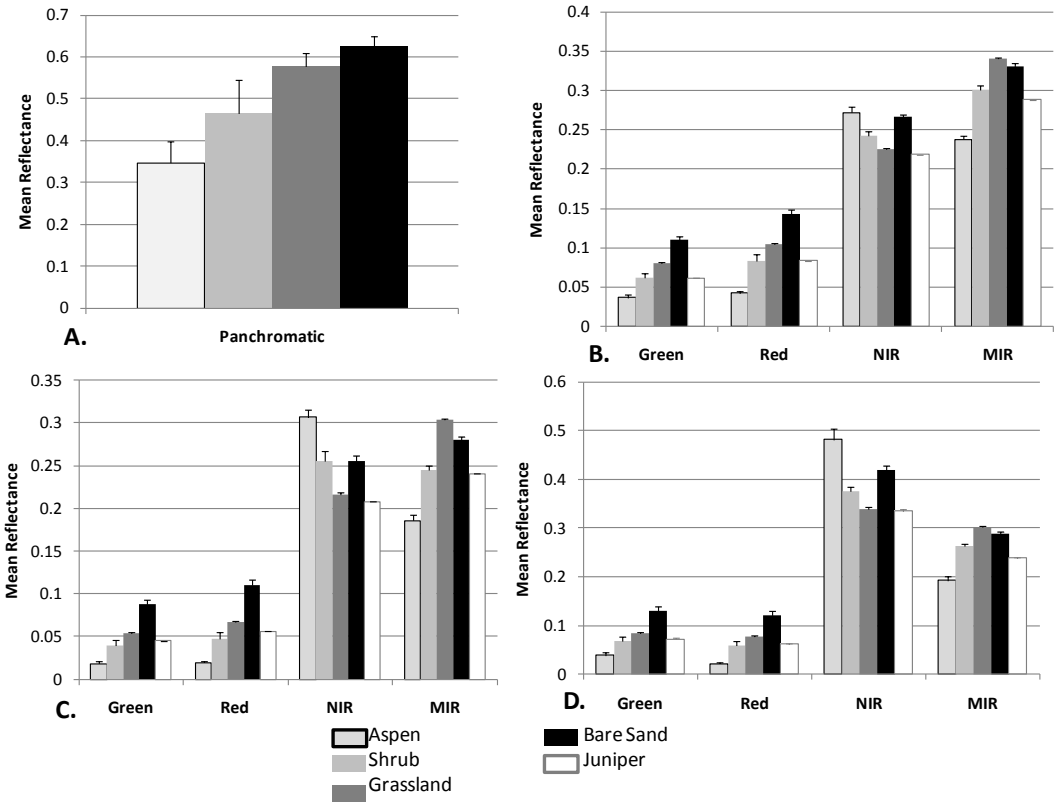


Figure 2.4: Spectral curves of the land cover classes within SPOT5 imagery. Figure A. shows the average reflectance of each land cover class within the SPOT5 panchromatic image from April 2007. Figure B. shows the average reflectance of each land cover class within the SPOT5 multispectral image from May 2009. Figure C. shows the average reflectance of each land cover class within the SPOT5 multispectral image from July 2009. Figure D. shows the average reflectance of each land cover class within the SPOT5 multispectral image from August 2007.

2.3.7 Accuracy Assessment

Accuracy assessment was completed using the fourteen validation data points collected in the field for each class to identify fourteen objects for each class. A confusion matrix was generated to compare predicted classes to actual classes based on the number of objects (Platt and Rapoza, 2008). The confusion matrix is defined in Table 2.1. Classification accuracy was evaluated from producer's and user's accuracies to give a measure of omission and commission error respectively. Producer's accuracy was calculated as agreement over the number of validation data objects in a class as defined in Table 2.1 (Congalton, 1991). User's accuracy was calculated as agreement over the number of classified objects in a class as defined in Table 2.1 (Congalton, 1991). Overall accuracy was calculated as the sum of agreement over the total number of validation data objects as defined in Table 2.1 (Congalton, 1991). Kappa statistics were also calculated to give a measure of agreement. Kappa values were calculated from the confusion matrices as

$$(Observed\ Agreement - Chance\ Agreement)/(1 - Chance\ Agreement) \quad (2.4)$$

where observed agreement is defined as overall accuracy (Naesset, 1996). Chance agreement was defined as

$$\sum i\ [\left(\frac{n}{N}\right)\left(\frac{n*}{N}\right)] \quad (2.5)$$

where *i* is the summation of each class, *n* is the number of classified objects in a class, *n** is the number of validation data objects in a class, and *N* is the total number of validation data objects as defined in Table 2.1 (Naesset, 1996). A minimum kappa value of 0.60 was deemed an acceptable level of accuracy because it represents substantial agreement with ground referenced data (Landis and Koch, 1977).

Table 2.1: Confusion matrix based on fourteen validation data objects per class. Data values are number of objects. (N) is the total number of validation data objects used in the accuracy assessment, (n*) is the number of validation data objects in a class and (n) is the number of classified objects in a class.

Class	Aspen	Shrub	Juniper	Grassland	Bare Sand	Sum (n)
Aspen	11	1	0	0	0	12
Shrub	3	11	3	0	0	17
Juniper	0	0	8	0	0	8
Grassland	0	2	2	13	2	19
Bare Sand	0	0	1	1	12	14
Sum (n*)	14	14	14	14	14	70 (N)

2.3.8 Class Metrics

To define the spatial relationships between land cover classes, class scaled metrics were calculated using Fragstats software (McGarigal and Marks, 1995). The output land cover classified map of the study site was converted to an ASCII text file and imported into Fragstats. Classes were numbered where 1 was aspen, 2 was shrub, 3 was juniper, 4 was grassland, and 5 was bare sand. Due to issues of correlation among metrics, metrics can be divided into six major categories which are able to represent ecological parameters of interest: patch area, patch shape, core area, isolation/connectivity, edge, and aggregation (Wang and Malanson, 2007). Other studies have found specific patch metrics such as number of patches, patch area, edge-density, and mean shape index were best at explaining habitat features for plant species (Uuemaa et al., 2009). Several metrics including percentage of landscape, number of patches, largest patch

index, mean patch size, patch size coefficient of variance, and landcape shape index were calculated because they are useful descriptors of landscape and class structure and they are non-redundant in the information they present (Herrera et al., 2009). A meaningful interpretation of landscape pattern should include information on the area, distribution, and shape of different land cover classes (Herrera et al., 2009).

The number of patches (NP) in each class were counted as a measure of class occurrence within the study site. The percentage of the landscape covered by each class was calculated as

$$\%Land = \frac{\sum_{j=1}^{n} aij}{A} (100) \tag{2.6}$$

where *aij* is the sum of the areas of all patches of the corresponding patch type and *A* is the total landscape area (McGarigal and Marks, 1995). Percent landscape gives a measure of the rarity of a certain class in the study site. The patch density was calculated as

$$PD = \frac{ni}{A} (100) \tag{2.7}$$

where *ni* is the number of patches of the corresponding patch type and *A* is the total landscape area (McGarigal and Marks, 1995). It is a measure of the number of patches per 100 m². The average area of a patch within a class was calculated as

$$MPS = \frac{\sum_{j=1}^{n} aij}{ni} \tag{2.8}$$

where *aij* is the sum of the areas of all patches of the corresponding patch type and *ni* is the number of patches of the same type (McGarigal and Marks, 1995). The measure is given in units of square meters. The area coefficient of variance was calculated as

$$PSCV = \frac{PSSD}{MPS} \tag{2.9}$$

defined as standard deviation divided by the mean (McGarigal and Marks, 1995). The PSCV is a measure of the relative variability in patch size for each class. Cohesion gives a quantitative measure of the connectivity of a class by calculating the perimeter to area ratio and dividing it by a patch shape index (McGarigal and Marks, 1995). The landscape shape index is defined as

$$LSI = \frac{0.25 \sum_{k=1}^{m} eik}{\sqrt{A}} \tag{2.10}$$

where it is the sum of all edge segments involving the corresponding patch type divided by the square root of the total landscape area, adjusted by a constant for the square standard in raster data (McGarigal and Marks, 1995). LSI will equal one if patches are all square and will increase from one as shape becomes more irregular or the length of edges increase. The largest patch

index gives a measure of the percentage of the area of a class taken up by the largest patch in that class where LPI approaches zero when the largest patch is increasingly small (McGarigal and Marks, 1995). It can be written as

$$LPI = \frac{\max aij}{A} (100) \quad (2.11)$$

where *max aij* is the area of the largest patch in the corresponding class and *A* is the total area of the class (McGarigal and Marks, 1995). The contrast weighted edge density was calculated as

$$CWED = \frac{\sum_{k=1}^{m'}(eik*dik)}{A} \quad (2.12)$$

where *(eik*dik)* is the sum of the lengths of each edge segment of the corresponding patch type multiplied by the corresponding contrast weight and *A* is the total landscape area (McGarigal and Marks, 1995). The measurement is in units of meters per square meter. The CWED increases as the amount of edge in a class increases or as the contrast in edges increases (McGarigal and Marks, 1995). The contrast weighted matrix for defining the contrast weights is presented in Table 2.2 where a value of one signifies complete contrast between classes and a value of zero represents no contrast between classes.

Table 2.2: Contrast weights between land cover classes for use in calculating the contrast weighted edge density. A value of one signifies complete contrast between classes and a value of zero signifies no contrast between classes.

CW Matrix	Aspen	Shrub	Juniper	Grassland	Bare Sand	Other
Aspen	0	0.2	0.8	0.9	1	1
Shrub	0.2	0	0.1	0.5	0.8	1
Juniper	0.8	0.1	0	0.35	0.5	1
Grassland	0.9	0.5	0.35	0	0.2	1
Bare Sand	1	0.8	0.5	0.2	0	1
Other	1	1	1	1	1	0

2.4 Results and Discussion

2.4.1 Scales in Landscape Patterns and Remote Sensing Products

2.4.1.1 Temporal Resolution

2.4.1.1.1 Spectral Variability

Since the coefficient of variance (CV) is a measure of relative variance, it can be thought of as the overall phenological variation within a class (Krishnaswany et al., 2004), as represented by the variation in spectral reflectivity between the twenty points for each class. Identical vegetation types can vary in reflectance values due to differences in coverage or vigour (Geerken et al., 2005), increasing the phenological variation within a class. In land cover classification, minimizing within class heterogeneity will increase between class separability. If internal heterogeneity is too high, threshold values in the classification will overlap, increasing confusion in the classification (Ju et al., 2005).

The CV determined that the shrub class had the most internal heterogeneity in the early growing season (Table 2.3). The shrub class consists of choke cherry, snowberry, and wolf willow, all of which exhibit different phenological cycles, thus increasing the variability in reflectance within this class. Variation within the shrub class was lowest in the late growing season indicating that the different shrub species were more closely matched in their phenological stage at this time of year as compared to the early growing season. The MIR and green bands showed the most internal homogeneity for the shrub class. In contrast, the evergreen shrub, juniper, showed the lowest amount of internal heterogeneity in June in the NIR band (Table 2.3). Grassland exhibited the most internal homogeneity in the middle growing season. Internal variation was lowest in the MIR and NIR bands for the grassland class signifying homogeneity in vegetation vigour/coverage. Variation within the bare sand class was lowest in the late growing season (Table 2.3). Bare sand within the study site exhibits slight to partial vegetation coverage, increasing the variability in spectral reflectance within that class. This variability can be exaggerated or minimized as the growing season progresses due to the variability in the phenological stages of the vegetation present on the sand (Lucas et al., 2007). In general, the most spectral homogeneity within classes was observed in the late growing season (Table 2.3).

Table 2.3: Coefficient of Variance (CV) for selected classes and bands across the growing season of 2009. A smaller CV value represents more homogeneity within classes. Generally, August data showed the most internal homogeneity.

Class	CV blue	CV green	CV red	CV NIR	CV MIR
June					
Shrub	0.51	0.45	0.48	0.45	0.34
Grassland	0.15	0.17	0.18	0.11	0.08
Juniper	0.30	0.13	0.19	0.12	0.31
Bare Sand	0.17	0.22	0.26	0.23	0.32
July					
Shrub	0.39	0.28	0.34	0.27	0.20
Grassland	0.14	0.13	0.13	0.13	0.12
Juniper	0.29	0.17	0.24	0.14	0.28
Bare Sand	0.20	0.21	0.22	0.17	0.13
August					
Shrub	0.40	0.26	0.30	0.27	0.26
Grassland	0.19	0.17	0.18	0.14	0.14
Juniper	0.25	0.14	0.19	0.13	0.23
Bare Sand	0.17	0.20	0.20	0.13	0.11

2.4.1.1.2 Spectral Separability

In the early growing season, the shrub class exhibited the highest reflectance in the NIR band signifying the most coverage of green vegetation as compared to other classes (Figure 2.5). Comparatively, bare sand exhibited the highest reflectance in the red band because, unlike vegetation, bare soil does not absorb red light (Zhang and Guo, 2008). Bare sand exhibited little to no change in spectral reflectance across the growing season (Figure 2.5-2.7). This is typical of areas with sparse vegetation cover dominated by bare soil because land cover is temporally uniform showing little intra-annual dynamics (Geerken et al., 2005). In the early growing season, grassland showed higher reflectance in the green and red bands than throughout the rest of the growing season. It has been identified that spectral reflectivity within the mixed-grass prairies is highly affected by standing dead materials, litter, biological crust, and bare soil due to the sparse coverage of grasses in these areas (Zhang and Guo, 2007). These variables contribute to increased reflectance of green and red light, particularly in the early growing season when coverage of grass species is low. As green-up occurs in grass species, the higher coverage of green vegetation absorbs more incident and red light causing reflectance in these bands to decrease as the maximum growing season is approached (Zhang and Guo, 2008).

Reflectance in the NIR band showed the greatest change from June to July following initial green-up (Figure 2.5-2.6). A greater change was observed in the shrub class than the juniper class because juniper is an evergreen, remaining relatively temporally uniform compared to the shrub class. As well, the largest change in NDVI occurred from June to July, signifying an increase in green vegetation during this time period (Figure 2.8). NDVI is most sensitive to changes in fractional vegetation cover, experienced during green-up in the early to middle growing season, as compared to changes in full vegetation cover (Carlson and Ripley, 1997). Shrub and grassland exhibited the largest increase in NDVI from June to July compared to juniper and bare sand. Bare sand characteristically exhibited no change in NDVI over the growing season due to the sparse vegetation cover and the over riding influence of bare soil on the spectral signature (Jensen, 2005).

Spectral separability between classes was lowest in the middle growing season (Figure 2.6). However, bare sand could be spectrally separated based on the high reflectance in the red band and grasslands stood out in the MIR band. The MIR band is sensitive to moisture such that as the moisture content of vegetation increases, reflectance in the MIR band decreases because water is a good absorber of MIR light (Jensen, 2005). Precipitation and soil water-holding capacity are one of the limiting factors controlling the growth of grasses in the mixed-grass prairie (Sala et al., 1988). Total precipitation in the middle growing season reached a low at 58.5 mm (July 2009) compared to 81.5 mm in the early growing season (June 2009) and 90.5 mm in the late growing season (August 2009) (Environment Canada, 2009). The relatively drier conditions in the middle growing season are seen in the MIR band of Figure 2.6. Bare sand and grassland had higher reflectance in the MIR band, while shrub and juniper had a lower reflectance in the MIR band. The low soil water-holding capacity of sandy soils allows water to percolate deeper into the ground, decreasing water loss from surface runoff and evapotranspiration (Sala et al., 1988). Differences in the rooting depths between species can account for differences in water uptake from sandy soils during dry periods. Species that are more permanent in dry sandy habitats tend to be deeply rooted while other species will develop extensive mats of shallow roots and rhizomes which more efficiently utilize surface moisture (Hulett et al., 1966). Grass species tend to rely more on surface soil water while shrub species tend to rely more on deep soil water (Nippert and Knapp, 2007).

In the late growing season, bare sand and juniper stood out with a high reflectance in the red band and a low reflectance in the MIR band, respectively (Figure 2.7). Juniper showed a slight decrease in reflectance in the NIR band across the growing season but remained almost unchanged in the other bands (Figure 2.5-2.7). Considering juniper is an evergreen shrub, it does not follow the same phenological cycle typical of other shrubs in the study site. Therefore, a spike in reflectance in the early growing season is not observed in its' spectral response curve because it does not undergo green-up in the same way. The slight differences in reflectance across the growing season could be influenced by abiotic factors such as drought stress or increased rainfall periods (Geerken et al., 2005). It was noted that in 2009, the early growing season was relatively cool and wet (average temperature of 14.91ºC and total precipitation of 81.5 mm in June), with a slight increase in temperature in the middle and late growing seasons (Environment Canada, 2009). Precipitation was relatively low in the middle growing season but spiked in the late growing season at 90.5 mm (Environment Canada, 2009). The decrease in NIR reflectance over the growing season signifies a decrease in vegetation vigour perhaps influenced by the relatively cool, wet autumn experienced in 2009.

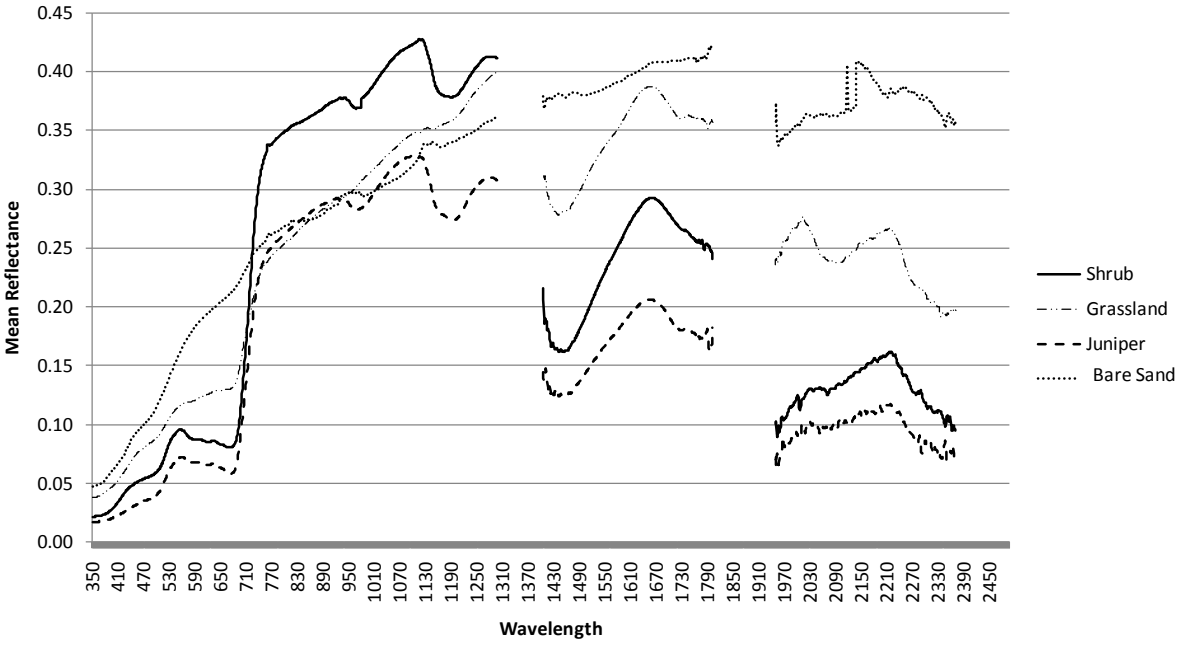


Figure 2.5: Spectral response curves of selected classes in the early growing season (June 2009).

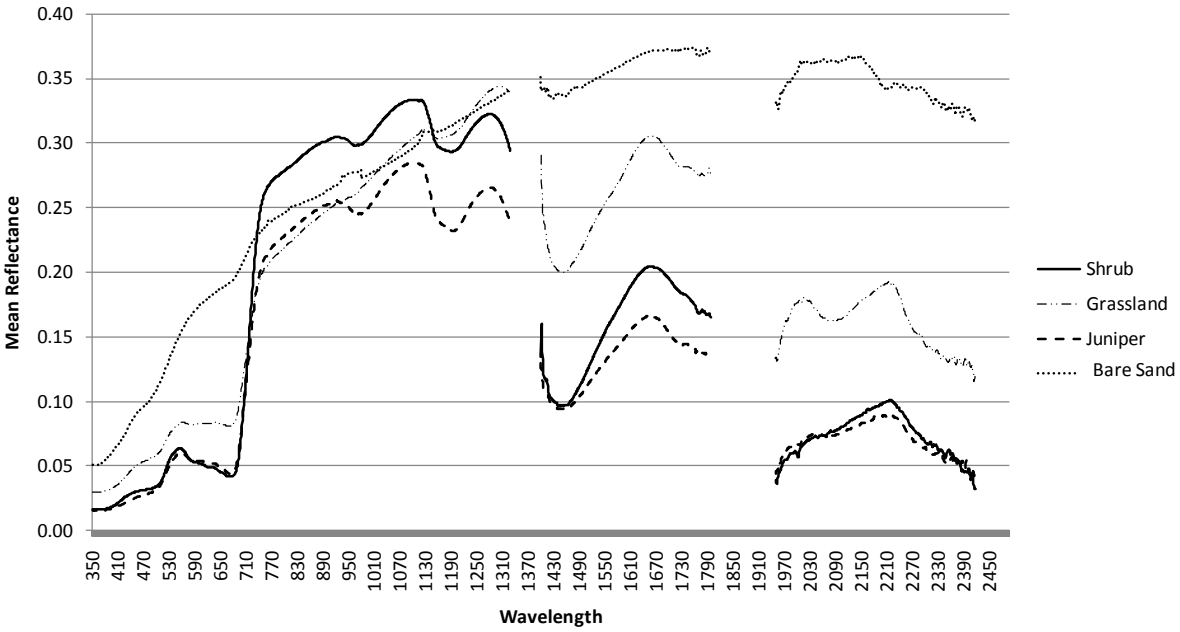


Figure 2.6: Spectral response curves of selected classes in the middle growing season (July 2009).

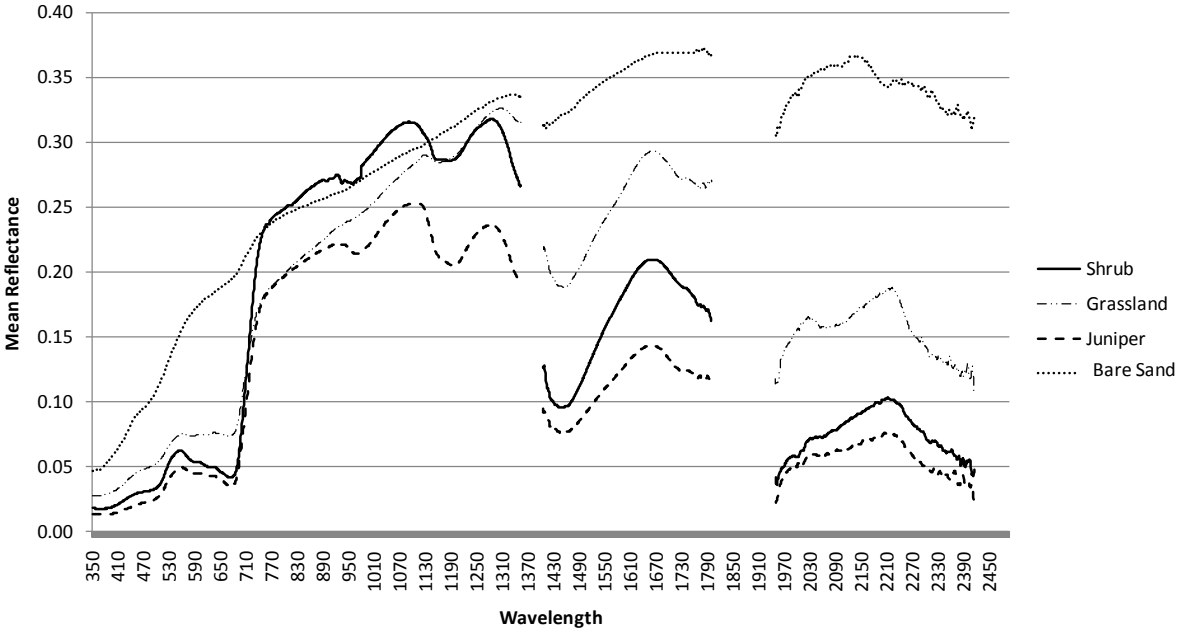


Figure 2.7: Spectral response curves of selected classes in the late growing season (August 2009).

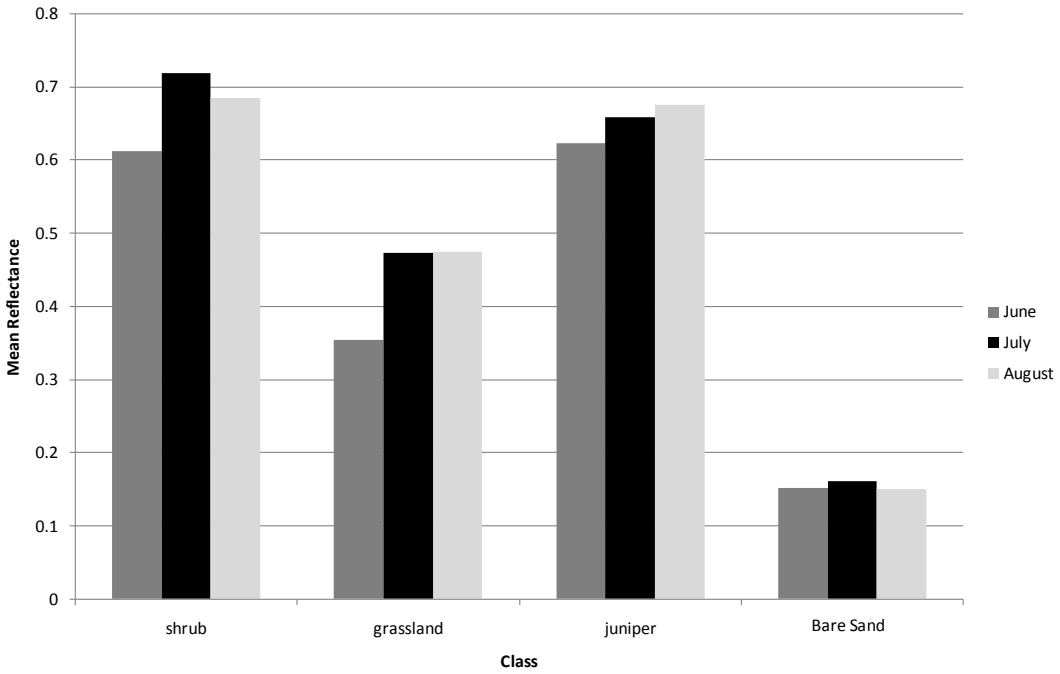


Figure 2.8: Average reflectance in NDVI for selected classes in the early, middle, and late growing seasons of 2009.

Different types of vegetation will absorb, reflect, and transmit light differently throughout the growing season, therefore image acquisition should occur at several times when contrasts in the spectral response curves are exaggerated enough to differentiate between them (Jin et al., 2007; Lucas et al., 2007). The spectral bands that can differentiate the different land cover classes at different times during the growing season are shown in Figure 2.9. In order to capture all four land cover classes, reflectance in the red and MIR bands in the late growing season, and green and NIR bands in the early growing season are needed at minimum (Figure 2.9). Therefore, imagery from the early and late growing season would be most useful for land cover classification of this study site. The panchromatic and red bands, with characteristically high reflectance being equated to bare sand, are most useful for bare sand classification. Similar to the results found here, Lucas et al. (2007) found that imagery acquired in the early growing season during leaf flush was needed to discriminate between deciduous and evergreen species, while imagery from the late growing season highlighted differences between grasslands and upland communities for agricultural land cover mapping.

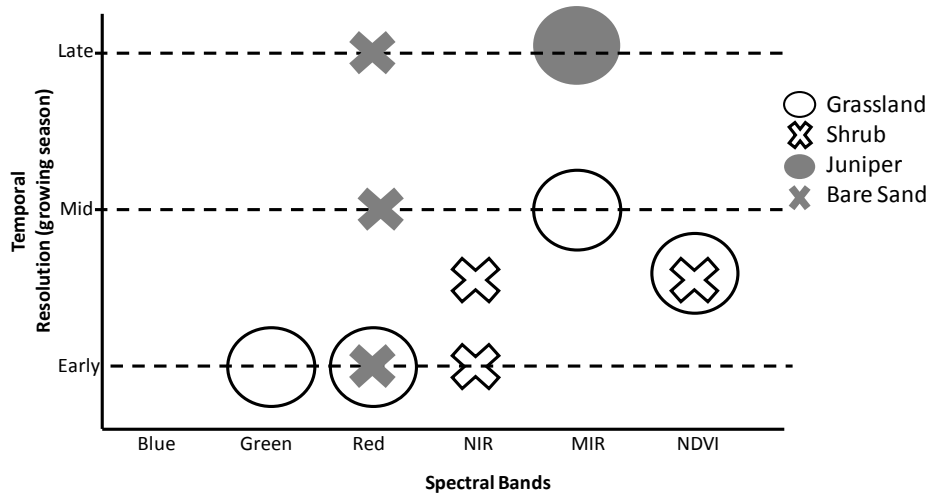


Figure 2.9: Spectral bands that are useful for discriminating between land cover classes at different times during the growing season. Band width is equivalent to SPOT5 multispectral imagery.

2.4.1.2 Spatial Resolution

Wavelet analyses for wavelengths 490-690 nm in the late growing season identified significant spatial scales between 8-20 m and around 64 m (Figure 2.10). The wavelet spectrum of the field transect for wavelengths 610-680 nm shows similar spatial scales at 8-16 m and around 64 m (Figure 2.11). Due to the heterogeneity in grassland composition, productivity, and diversity, the spatial patterns observed within grassland ecosystems operate at different spatial scales (He et al., 2007). Therefore, different biological variables should be studied at different spatial scales (Zhang and Guo, 2007). For example, He et al. (2007) found that topography varied at a spatial scale of 120 m in Grasslands National Park while soil moisture varied at scales around 20 m. The spatial scales identified in this study are smaller than those found in Grasslands National Park which can be attributed to geographical differences in the study sites, such as different soil texture and chemistry, and different topographical gradients. However, as found in Grasslands National Park, two levels of spatial variation were identified.

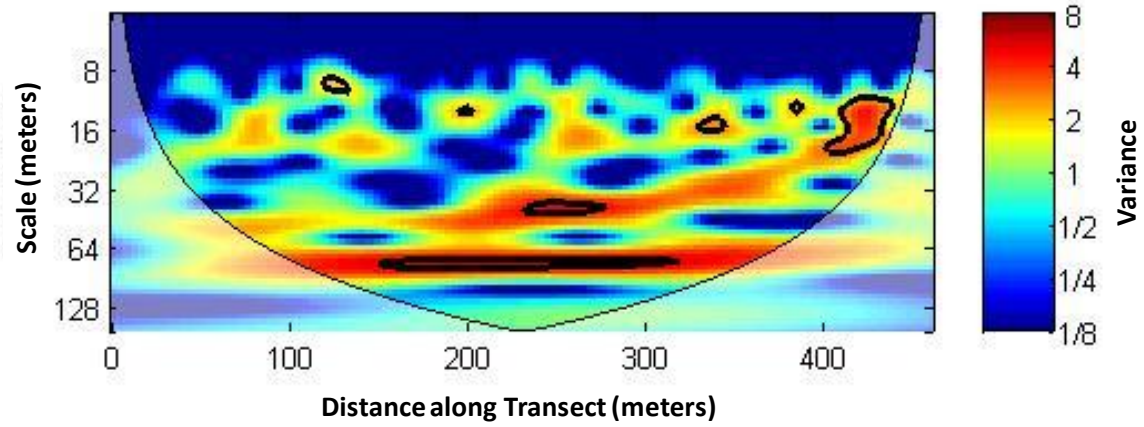


Figure 2.10: Wavelet spectrum of the field transect in the late growing season of 2009 for the panchromatic band. The morlet wavelet mother function was used to analyze the panchromatic band (average reflectance of wavelengths 490-690 nm). The transect consisted of 128 points spaced 4 m apart. Significant spatial scales are circled in black.

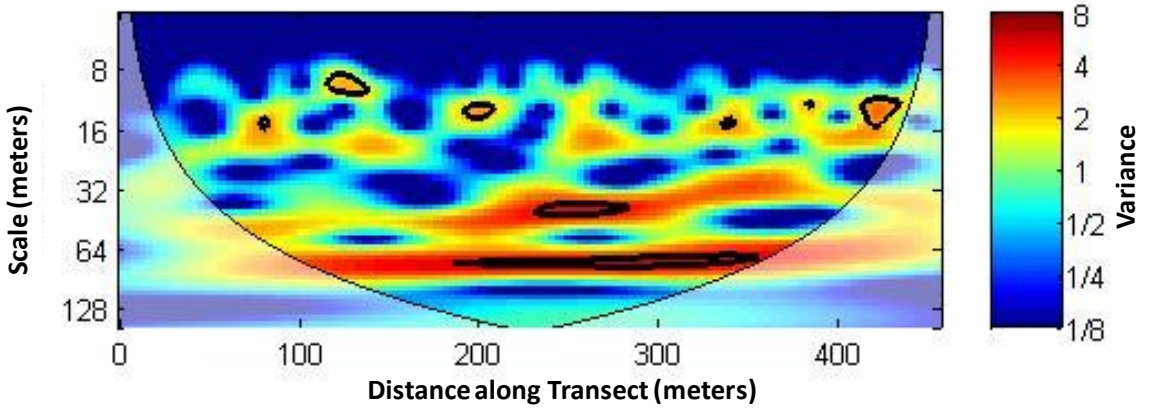


Figure 2.11: Wavelet spectrum of the field transect in the late growing season of 2009 for the red band. The morlet wavelet mother function was used to analyze the red band (average reflectance of wavelengths 610-680 nm). The transect consisted of 128 points spaced 4 m apart. Significant spatial scales are circled in black.

Significant spatial variation at scales between 8-20 m was observed at around 140-144 m, 212 m, 360-364 m, and 428 m along the transect (Figure 2.12). These transect locations are associated with the occurrences of bare sand patches. At 140 m along the transect, about 83% of the cover was bare sand and 23% of the cover was bare sand at 144 m along the transect. At 212 m about 53% of the cover was bare sand. At 360 m about 44% of the cover was bare sand with bare sand cover increasing to about 63% at 364 m. At 428 m about 25% of the cover was bare sand. Therefore, following the sampling theorem, the optimum spatial resolution for detecting

bare sand would be 2-5 m (1/4 of the wavelet scale at 8-20 m) (Guo et al., 2001; Rahman et al., 2003).

The spatial pattern of grassland ecosystems is scale dependent. Large-scale patterns in grasslands, such as noted at 64 m in the wavelet spectrum, are determined by variations in topography or climate. Small-scale pattern is influenced from abiotic and biotic factors such as soil composition or grazing pressure (He et al., 2007). Topography within the study site is influenced by the existence of dune complexes oriented in a general north-west to south-east direction. This low dunes ecosite has a general topography of low-lying hills interspersed in terrain that is almost flat (Thorpe, 2007). Slope aspect and slope position highly influence local plant communities within the study site, causing small-scale patterns as noted at 8-20 m in the wavelet spectrum. Ridge tops and south facing slopes tend to have less vegetation and more bare sand with north facing slopes typically vegetated by woody species. South-facing slopes tend to lack the development of an A-horizon, confining the plant community to such species as sand reed grass. However, vegetation on south-facing slopes tends to increase as the slope gradient decreases because of greater soil stability (Thorpe, 2007). The interaction of these localized abiotic factors influences small-scale patterns within the landscape. Therefore, using coarse spatial resolution imagery may miss the small-scale patterns that are ecologically significant in a prairie sand dune landscape (He et al., 2007).

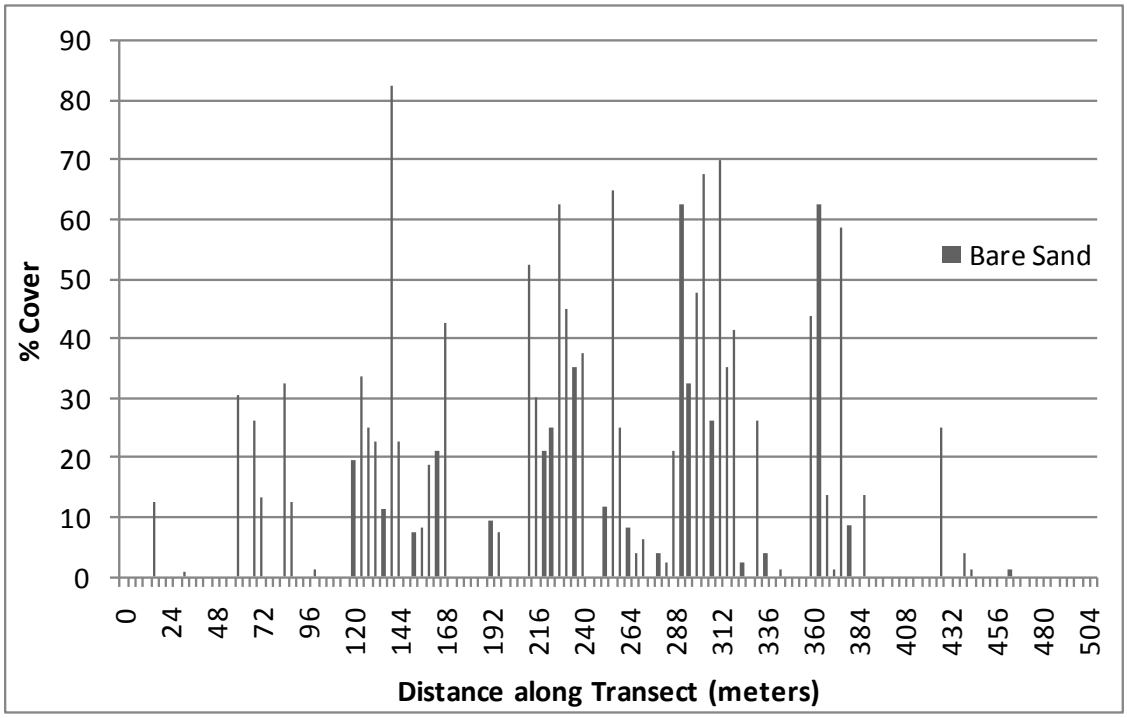


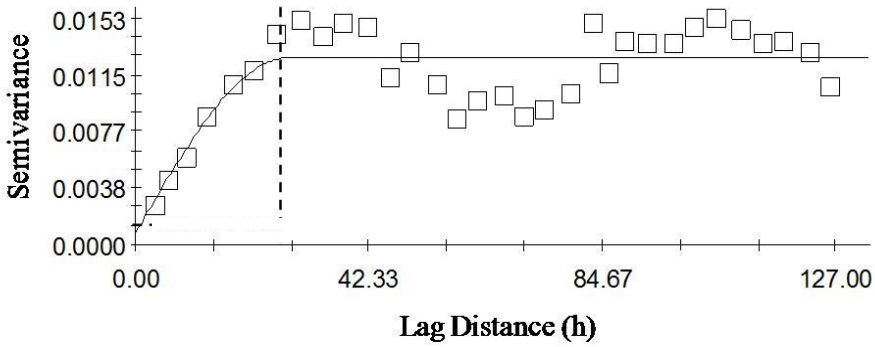
Figure 2.12: Percent bare sand along the field transect as corresponding with significant locations in the wavelet spectrum (Figure 2.10 and Figure 2.11). Significant spatial variations observed in the wavelet spectrum correspond to bare sand occurrences at 140-144 m, 212 m, 360-364 m, and 428 m along the transect.

2.4.1.3 Effect of Scale Parameter on Image Segmentation

Semivariogram analysis revealed a lag distance of about 28 m beyond which pixels in the panchromatic image were no longer spatially autocorrelated (Figure 2.13). Taking one half of the semivariogram lag distance, the scale parameter in image segmentation should be adjusted such that the average distance between neighbouring image object centroids is around 28 m (McGrew and Monroe, 2000; Karl and Maurer, 2010). At a scale parameter of ten, the average distance between neighbouring image object centroids is 21.2 m with a standard deviation of 7.7 m, near to the lag distance identified in the semivariogram (Figure 2.14). At a scale parameter of fifteen and twenty, the average distance between neighbouring image object centroids exceeds the semivariogram lag distance (Figure 2.14). At these scale parameters, spectral heterogeneity within objects is exceeding that which is statistically significant and pixels which are not spatially autocorrelated are being grouped together into objects. At a scale parameter of five, the average distance between neighbouring image object centroids is well below the lag distance of 28 m. Objects at this scale parameter are still spatially autocorrelated and could be merged to reduce noise in the classification.

Over segmenting an image can create artificial pattern that does not occur on the ground while under segmenting an image can cause the loss of pattern that does occur on the ground. The selection of an optimal spatial resolution and scale parameter is desirable in order to preserve the boundaries between classes observed on the ground and to produce an ecologically relevant land cover classification (Ju et al., 2005; Moller et al., 2007). During image segmentation, optimal parameter settings occur when under and over segmentation are balanced (Moller et al., 2007). In this study, over and under segmentation were balanced at a scale parameter of ten as seen in the desirable lag distance between neighbouring image object centroids. At scale parameters greater than ten, under segmentation was observed since the resolution of image objects was exceeding the desirable lag distance. In this case, the boundaries between land cover types were not well defined causing misrepresentation of the study site.

A scaling threshold was reached at a lag distance of 28 m beyond which pixels were no longer spatially autocorrelated. This was represented in the segmentation by the size of image objects as controlled by the scale parameter. In this study, the scaling threshold was reached at a scale parameter of ten, beyond which spectral heterogeneity within objects was too high. Similar to the results noted here, Karl and Maurer (2010) found that semivariograms constructed from their field data indicated a scaling threshold/lag distance that was almost identical to the average distance between neighbouring image object centroids at the segmentation scale parameter beyond which field variables were no longer able to be correlated with image products. Ju et al. (2005) also found that coarsening the spatial scale initially improved their classification accuracy compared to pixel-based classification, but that accuracy quickly dropped off after a scaling threshold was reached. Therefore, an optimal scale parameter should preserve the boundaries between patches at finer scales to avoid under segmentation and smooth out the variation within patches at coarser scales to avoid over segmentation (Ju et al., 2005). In this study, the optimal scale parameter for image object segmentation was ten.



Spherical model (Co = 0.00077; Co + C = 0.01264; Ao = 27.60; r2 = 0.616; RSS = 1.256E-04)

Figure 2.13: Semivariogram for the SPOT5 panchromatic image transect fitted to a spherical model. The squares represent the data points, the solid line represents the model fit, and the dashed line represents the lag distance at 28 m.

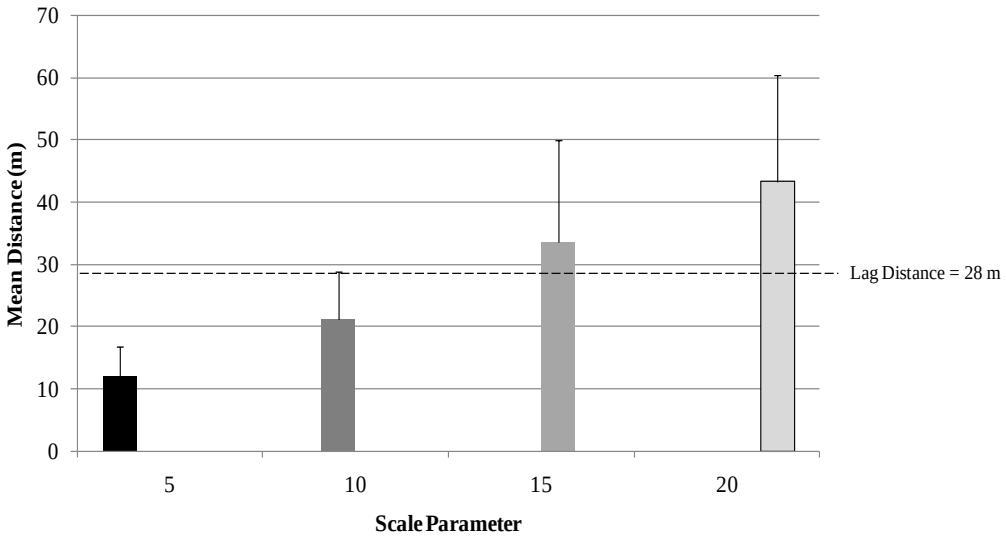


Figure 2.14: Average distance between neighbouring image object centroids across segmentations of the SPOT5 panchromatic image. Scale parameter was varied at five, ten, fifteen, and twenty with other parameters held constant. The lag distance identified in the semivariogram at 28 m is represented by the dashed line.

2.4.2 Land Cover Classification

Figure 2.15 shows the land cover classification of the study site. Overall classification accuracy was 79% and showed substantial agreement with ground referenced data (overall kappa

substantially higher than the minimum standard of 0.60) (Table 2.4). Herrera et al. (2009), Lucas et al. (2007), and Platt and Rapoza (2008) obtained similar overall accuracies in their land cover classifications using multi-temporal and/or multi-resolution data sets. Overall classification accuracy is affected by several problems inherent in grassland classification: compared to trees or shrubs, grass species are much smaller than any spatial resolution currently available; unlike crops, spatial variation in coverage will occur within a certain grassland type; and spatial variation in the types of species present within a certain grassland type will occur (Herrera et al., 2009). Grassland classification is also complicated by the existence of standing dead materials, litter, biological crust, and bare soil affecting the spectral signature of land cover classes (Zhang and Guo, 2007). Therefore, in classifications of mixed-grass prairie regions, it is rare to find pixels that are "pure" (Herrera et al., 2009).

Grassland and bare sand had the highest classification accuracy as represented by kappa values of 0.90 and 0.82 respectively (Table 2.4). However, bare sand were most reliably classified as seen in producer's and user's accuracies of 86%. The confusion matrix in Table 2.1 shows that while omission error was only 14%, bare sand was most often misclassified as grassland. Similarly, commission error at 14% showed that juniper and grassland were sometimes misclassified as bare sand (Table 2.1). High classification accuracy in bare sand can be attributed to spectral separability in the red and panchromatic bands (Figure 2.4). Although grassland obtained a higher kappa value than bare sand, reliability of grassland classification was lower as seen in a user's accuracy of 68%. Grasslands were over classified in the study site due to confusion with several classes, namely bare sand, juniper, and shrub (Table 2.1).

Juniper had an unacceptable level of classification accuracy with a kappa value below 0.60 (Table 2.4). This contributed to a decrease in the overall accuracy of the classification. Low classification accuracy in juniper can be attributed to the high degree of omission error at 43%. Juniper is under classified because it is being omitted from the classification due to confusion with shrub (Table 2.1). Confusion was highest between juniper and shrub due to the low spectral separability between them, as seen in Figure 2.4. Spectral separability of juniper within the multispectral bands was not possible (Figure 2.4) thus classification rules were based on reflectance in the panchromatic band. Juniper was first extracted as part of the shrub class and then separated from the shrub class based on the uniformity of patch shape. Therefore, confusion between shrub and juniper was higher than expected. Shrub and aspen also showed substantial

agreement with ground referenced data with kappa values of 0.72 and 0.74 respectively. The aspen class was somewhat under classified due to confusion with the shrub class (Table 2.1). Comparatively, the shrub class showed a higher degree of commission error as seen in a user's accuracy of 65%. Aspen and juniper were misclassified as shrub, contributing to an over classification of shrub in the study site (Table 2.1).

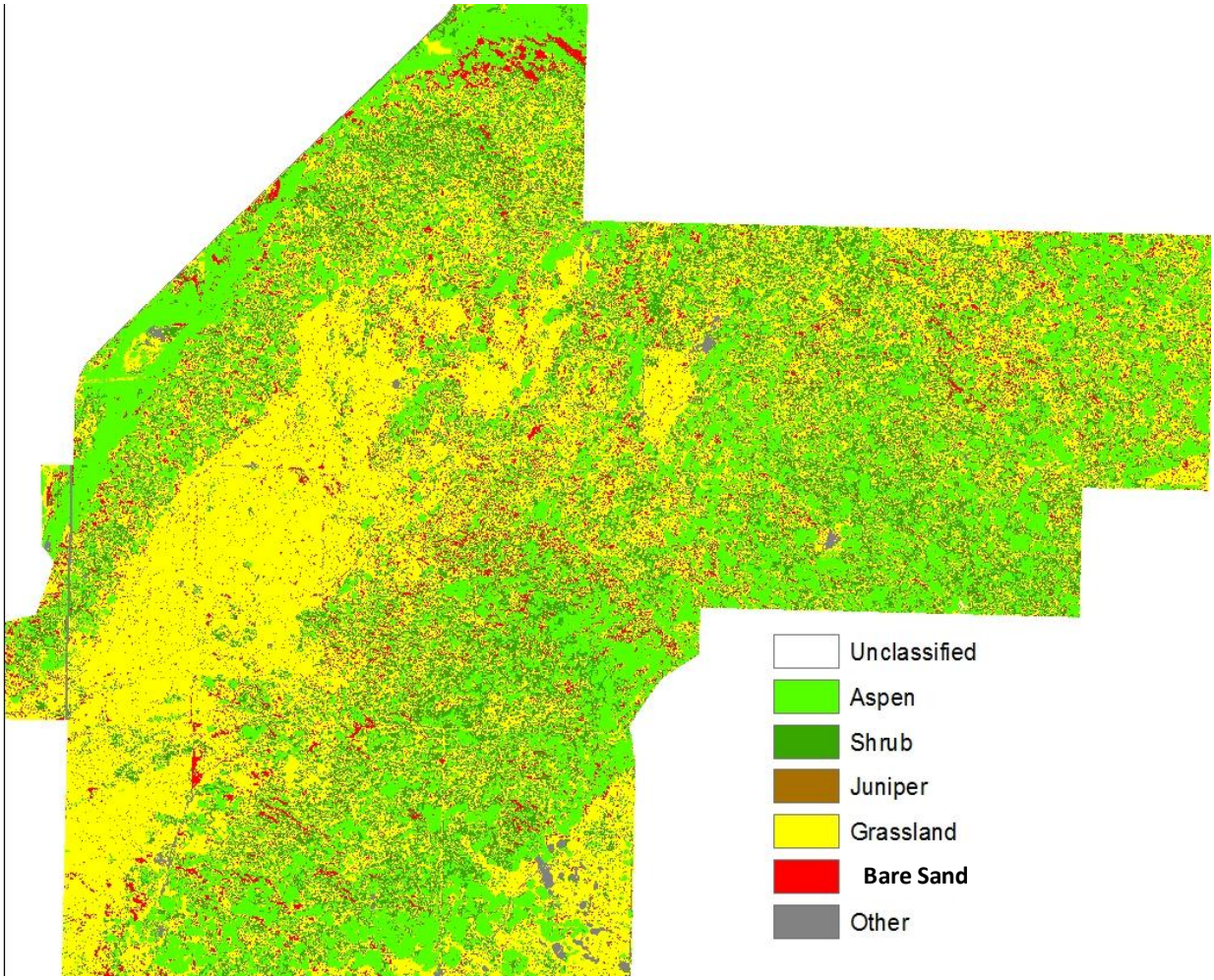


Figure 2.15: Land cover classification of the study site.

Table 2.4: Accuracy assessment for the land cover classification of the study site.

	Aspen	Shrub	Juniper	Grassland	Bare Sand
Producer (%)	79	79	57	93	86
User (%)	92	65	100	68	86
KIA Per Class	0.74	0.72	0.52	0.90	0.82
Overall Accuracy	79%				
Overall KIA	0.73				

2.4.3 Spatial Relationship between Land Cover Classes

Grassland comprised the matrix of the area covering 45.5% of the study site (Table 2.5). Aspen and shrub were the most dominating landscape elements comprising 25.5% and 19.2% of the study site respectively. Bare sand made up only 6.0% of the study site while juniper occupied the least amount of area comprising only 2.7% of the study site. The other 1.1% was masked out in the class "other", which included known water bodies, roads, trails, corrals, and other developements.

The juniper class had the most number of patches but, on average, had the smallest patch size and the least amount of cohesion between patches (Table 2.5). Creeping juniper is a groundcover species, with single plants growing in circular mats spreading from 3.5-6 meters (Miller, 1978). Considering the growth patterns of juniper, it is not surprising that this class is represented by a large amount of very small and individually dispersed patches on the landscape. The patch size coefficient of variance was relatively low for juniper at 0.94 compared to other classes, indicating that juniper plants tend to grow to similar sizes across the study site.

The matrix of a landscape can be identified because it is the most extensive and most connected landscape element (Forman, 1995). Just over 50% of the grassland class was comprised of one patch. As well, grassland exhibited the highest amount of cohesion compared to other classes. The grassland class showed the most irregularity in patch shape (Table 2.5). Considering grassland is the matrix of the study site, it will show high irregularity in patch shape due to concave borders created between landscape elements within the matrix (Forman, 1995). The highest contrast weighted edge density was observed in the grassland class because it encompasses all other land cover classes. The high variability in patch size, as indicated by the PSCV value, indicates that the distribution of grassland patch area is heterogeneous.

The aspen class comprises about one quarter of the study site yet it has the lowest number of patches as compared to other land cover classes. Considering that aspen was under classified (omission error of 21%), this gives a conservative measure (Table 2.4). On average, patch size in the aspen class was much larger than patch size in any of the other classes. However, variability in patch size was also greater than other land cover classes, except grassland. This is also supported by the relatively high cohesion of aspen patches. Considering that aspen mainly reproduces by suckering to form large stands of "clones" (Maini, 1960), the growth pattern of this species will highly affect the spatial pattern of this class on the landscape. Therefore, even though there are only a small number of aspen patches in the study site, they are among the largest and most cohesive observed due to the nature of reproduction in aspen. As well, the high variability in patch size could be a reflection of the variability in stand age and/or the variability in environmental conditions between sites as these factors can exert control over reproduction and thus the size of a stand of clones. The contrast weighted edge density for aspen is the next highest after grassland. The most contrast with aspen occurs with bare sand and grassland. Therefore, the higher contrast weighted edge density indicates that aspen is bordering more with bare sand and grassland than with shrub or juniper.

Comparatively, the shrub class (snowberry, choke cherry, wolf willow) makes up just under one quarter of the study site. The shrub class is composed of many small patches that are relatively cohesive and irregularly shaped (Table 2.5). The shrub class also has the third highest contrast weighted edge density, again signifying its border with grassland and sand dunes as opposed to aspen stands. Fire, bison grazing, and the interaction of the two processes have historically been essential to the maintenance of grassland structure and composition within the mixed-grass prairie region (Steuter et al., 1990). Fire and bison grazing will change above ground net primary production favouring increased growth of grass species, an essential role in natural succession (Steuter et al., 1990; Hobbs and Huenneke, 1992). Recent fire suppression and a change to cattle grazing in the study site may affect these successional trends, negatively influencing population growth rates of herbaceous species and allowing for the relatively high proportion of the study site to be occupied by woody species (Vitt et al., 2009).

Small patches are more affected by the nature of the matrix than larger patches, therefore the grassland matrix becomes a factor in determining the degree of isolation between small patches (Lord and Norton, 1990). However, small patches that more densely populate the matrix

can influence the nature of the intervening matrix (Lord and Norton, 1990). Other than juniper, bare sand exist as some of the smallest patches within the study site. However, the number of sand dune patches within the study site is less than half that of juniper. While bare sand comprise a larger percentage of the study site than juniper, juniper has a much higher patch density. Therefore, the influence of the matrix on landscape elements is highest in bare sand patches. Further, a relatively low PSCV value indicates a homogeneous distribution of bare sand size across the study site. The fairly even distribution of bare sand patch size is also seen in the low LPI value, where less than 1% of the bare sand class is comprised of a single large sized patch. Therefore, it can be assumed that matrix influence is equal on all bare sand patches. Bare sand exhibits a relatively low contrast weighted edge density and is somewhat cohesive. Cohesion may indicate that bare sand occurs as a series of complexes rather than as individually isolated patches within the study site. The lower contrast weighted edge density also indicates that bare sand is more likely to be bordered by grassland or juniper classes as opposed to shrub or aspen. Bordering grassland as opposed to shrub or aspen will increase bare sand cohesion and decrease isolation between patches because the gradient in boundary characteristics is less steep between bare sand and grassland allowing for better flow of materials, energy, and organisms between patches (Cadenasso et al., 2003).

Table 2.5: Class scaled metrics for the selected land cover classes from the study site classification. Where % LAND is the percentage of the landscape, NP is the number of patches, PD is patch density (number of patches per 100 square meters), MPS is the mean patch size (square meters), PSCV is the patch size coefficient of variation (unitless), CWED is the contrast weighted edge density (meters per square meter), COHESION (unitless), LPI is the largest patch index (%), and LSI is the landscape shape index (unitless).

Class	% LAND	NP	PD	MPS	PSCV	CWED	COHESION	LPI	LSI
Shrub	19.2	43027	0.038	503.7	4.81	0.0274	74.8	0.76	956.6
Aspen	25.5	11480	0.010	2506.8	13.51	0.0459	89.0	9.04	938.7
Juniper	2.7	73851	0.066	40.4	0.94	0.0053	50.2	0.08	492.7
Other	1.1	796	0.001	1561.9	3.81	0.0039	74.2	8.78	250.2
Grassland	45.5	32582	0.029	1574.0	97.56	0.0612	95.6	53.67	1063.1
BareSand	6.0	34505	0.031	195.0	3.41	0.0115	65.0	0.71	608.4

2.5 Conclusion

Understanding the appropriate application of scale in an ecological and technological context can traverse the technology-ecology gap. For this landscape, the optimum spatial resolution for bare

sand classification was between 2-5 m. Therefore, the data set for this land cover classification included a SPOT5 panchromatic image to take advantage of the high spatial resolution at 2.5 m. If data set resolution is not directly related to landscape spatial scale, remote sensing products may be either failing to detect pattern in a landscape where pattern actually exists or may be detecting manufactured pattern in a landscape where no pattern exists (Li and Wu, 2004). Spatial scale represents the magnitude of change that is occurring within landscape pattern as measured from ground based field surveys. This represents ecologically significant changes within the landscape. Wavelets or semivariograms can directly correlate spatial scale with image spatial resolution (Figure 2.16). Imagery spatial resolution should be one quarter of the wavelet scale or one half of the semivariogram scale (He et al., 2007) in order to accurately represent the spatial pattern observed on the ground in the spatial structure generated from the satellite image.

Intra-annual differences in vegetation phenology should also be considered in land cover classification. Image acquisition should occur several times throughout the growing season to better capture the spectral variability among land cover classes (Figure 2.16). In this study, it was found that the most appropriate temporal resolution for land cover classification and bare sand extraction in the mixed-grass prairie was early and late growing season. Therefore, three SPOT5 mulitspectral images from May, July, and August were selected to gain the needed temporal resolution. Spectral resolution obtained in the red, NIR, MIR, and panchromatic bands of the SPOT5 imagery was suitable for this land cover classification.

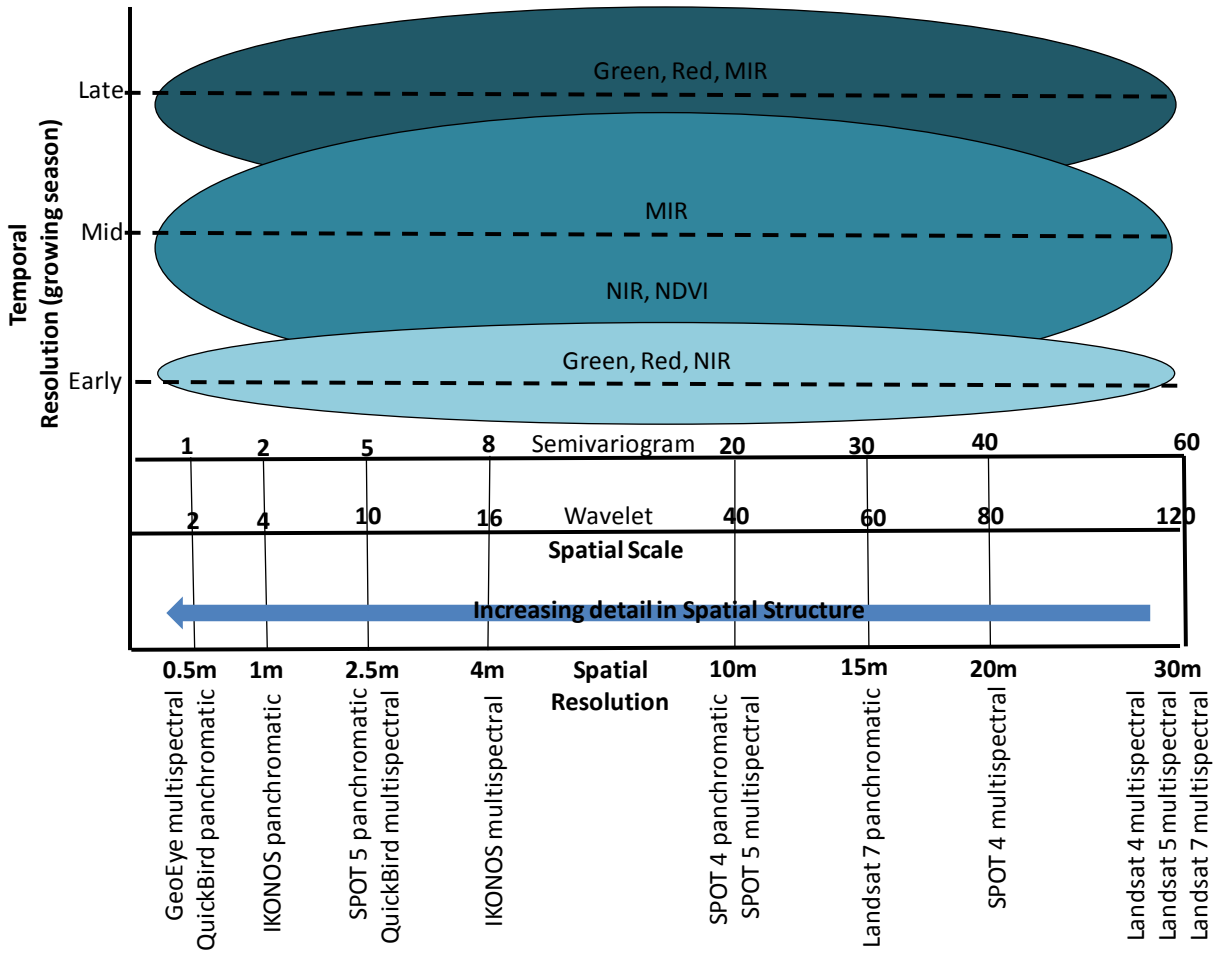


Figure 2.16: Relationship between ecological and technological scales of study.

Since the land cover classification in this study was conducted at scales that were significant in the landscape, the spatial relationships between land cover classes are ecologically meaningful. The spatial relationship between bare sand and other land cover classes appeared to be most influenced by the matrix of the study site. Bare sand patches exist as some of the smallest patches in the study site and are mainly situated within surrounding grassland areas. The grassland matrix will exert more influence over these smaller patches however, the magnitude of difference between these landscape elements is less steep allowing for increased flow of materials, energy, and organisms between patches (Cadenasso et al., 2003). Compositional gradients exist between landscape elements as a response to environmental gradients affecting the plant communities. Bare sand and sand dune structures tend to follow a compositional gradient in which natural succession moves towards dune stabilization and eventual grassland

communities (Hulett et al., 1966). For example, the gradient in the amount of light reaching the ground in a bare sand/grassland boundary is less steep than the change in light between a bare sand/aspen boundary (Cadenasso et al., 2003). Therefore, although structurally bare sand appears to be fragmented and isolated within the landscape, functional connectivity still remains intact. Therefore, bare sand patch characteristics cannot be understood in isolation from the surrounding grassland matrix.

2.6 References

Cadenasso, M., Pickett, S., Weathers, K., Jones, C., 2003. A framework for a theory of ecological boundaries. *BioScience* 53: 750-758.

Carlson, T. and Ripley, D., 1997. On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote Sensing of Environment* 62: 241-252.

Chen, K. and Blong, R., 2003. Identifying the characteristic scale of scene variation in fine spatial resolution imagery with wavelet transform-based sub-image statistics. *International Journal of Remote Sensing* 24: 1983-1989.

Congalton, R., 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment* 37: 35-46.

Elmqvist, B., Ardo, J., Olsson, L., 2008. Land use studies in drylands: an evaluation of object oriented classification of very high resolution panchromatic imagery. *International Journal of Remote Sensing* 29: 7129-7140.

Environment Canada, 2009. National climate data and information archive. URL: http://www.climate.weatheroffice.gc.ca/climateData/dailydata_e.html [Last accessed: Nov. 13, 2010].

Forman, R., 1995. *Land Mosaics: The Ecology of Landscapes and Regions.* Cambridge University Press, England.

Geerken, R., Zaitchik, B., Evans, J., 2005. Classifying rangeland vegetation type and coverage from NDVI time series using Fourier filtered cycle similarity. *International Journal of Remote Sensing* 26: 5535-5554.

Guo, X., Fu, B., Ma, K., Chen, L., 2001. Utility of semivariogram for spatial variation of soil nutrients and the robust analysis of semivariogram. *Journal of Environmental Sciences* 13: 453-458.

Gustafson, E., 1998. Quantifying landscape spatial pattern: What is the state of the art? *Ecosystems* 1: 143-156.

He, Y., Guo, X., Cheng Si, B., 2007. Detecting grassland spatial variation by a wavelet approach. *International Journal of Remote Sensing* 28: 1527-1545.

Herrera, L., Laterra, P., Maceira, N., Zelaya, K., Martinez, G., 2009. Fragmentation status of tall tussock grassland relicts in the flooding Pampa, Argentina. *Rangeland Ecology and Management* 62: 73-82.

Hirzel, A. and Guisan, A., 2002. Which is the optimal sampling strategy for habitat suitability modelling. *Ecological Modelling* 157: 331-341.

Hobbs, R. and Huenneke, L., 1992. Disturbance, diversity, and invasion: Implications for conservation. *Conservation Biology* 6: 324-334.

Hulett, G., Coupland, R., Dix, R., 1966. The vegetation of dune sand areas within the grassland region of Saskatchewan. *Canadian Journal of Botany* 44: 1307-1331.

Jensen, J., 2005. *Introductory Digital Image Processing 3rd Edition.* Prentice-Hall Inc., U.S.A.

Jin, Z., Tian, Q., Chen, J., Chen, M., 2007. Spatial scaling between leaf area index maps of different resolutions. *Journal of Environmental Management* 85: 628-637.

Jobin, B., Labrecque, S., Grenier, M., Falardeau, G., 2008. Object-based classification as an alternative approach to the traditional pixel-based classification to identify potential habitat of the Grasshopper Sparrow. *Environmental Management* 41: 20-31.

Ju, J., Gopal, S., Kolaczyk, E., 2005. On the choice of spatial and categorical scale in remote

sensing land cover classification. *Remote Sensing of Environment* 96: 62-77.
Karl, J. and Maurer, B., 2010. Multivariate correlations between imagery and field measurements across scales: comparing pixel aggregation and image segmentation. *Landscape Ecology* 25: 591-605.
Krishnaswamy, J., Kiran, M., Ganeshaiah, K., 2004. Tree model based eco-climate vegetation classification and fuzzy mapping in diverse tropical deciduous ecosystems using multi-season NDVI. *International Journal of Remote Sensing* 25: 1185-1205.
Labat, D., 2005. Recent advances in wavelet analysis: Part 1. A review of concepts. *Journal of Hydrology* 314: 275-288.
Landis, B. and Koch, G., 1977. The measurement of observer agreement for categorical data. *Biometrics* 33: 159-174.
Li, H. and Wu, J., 2004. Use and misuse of landscape indices. *Landscape Ecology* 19: 389-399.
Lightowlers, C., Nelson, T., Setton, E., Keller, C., 2008. Determining the spatial scale for analysing mobile measurements of air pollution. *Atmospheric Environment* 42: 5933-5937.
Lord, J. and Norton, D., 1990. Scale and the spatial concept of fragmentation. *Conservation Biology* 4: 197-202.
Lucas, R., Rowlands, A., Brown, A., Keyworth, S., Bunting, P., 2007. Rule-based classification of multi-temporal satellite imagery for habitat and agricultural land cover mapping. *Journal of Photogrammetry and Remote Sensing* 62: 165-185.
Maini, J., 1960. Invasion of grassland by Populus tremuloides in the northern great plains. *University of Saskatchewan.* Thesis.
McDermid, G., Franklin, S., LeDrew, E., 2005. Remote sensing for large area habitat mapping. *Progress in Physical Geography* 29: 449-474.
McGarigal, K. and Marks, B., 1995. FRAGSTATS: spatial pattern analysis program for quantifying landscape structure. Gen. Tech. Rep. PNW-GTR-351. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 122 p.
McGrew, J. and Monroe, C., 2000. *An Introduction to Statistical Problem Solving in Geography* 2nd *edition.* Boston: McGraw-Hill, pp. 110.
Miller, J., 1978. An ecological study of creeping juniper (Juniperus horizontalis Moench.) in Montana. *Montana State University*. Thesis.
Moller, M., Lymburner, L., Volk, M., 2007. The comparison index: A tool for assessing the accuracy of image segmentation. *International Journal of Applied Earth Observation and Geoinformation* 9: 311-321.
Naesset, E., 1996. Conditional tau coefficient for assessment of producer's accuracy of classified remotely sensed data. *Journal of Photogrammetry and Remote Sensing* 51: 91-98.
Nippert, J. and Knapp, A., 2007. Linking water uptake with rooting patterns in grassland species. *Oecologia* 153: 261-272.
Platt, R. and Rapoza, L., 2008. An evaluation of an object-oriented paradigm for land use/land cover classification. *Professional Geographer* 60: 87-100.
Rahman, A., Gamon, J., Sims, D., Schmidts, M., 2003. Optimum pixel size for hyperspectral studies of ecosystem function in southern California chaparral and grassland. *Remote Sensing of Environment* 84: 192-207.
Sala, O., Parton, W., Joyce, L., Lauenroth, W., 1988. Primary production of the central grassland region of the United States. *Ecology* 69: 40-45.
Smith, G. and Fuller, R., 2001. An integrated approach to land cover classification: an example in the Island of Jersey. *International Journal of Remote Sensing* 22: 3123-3142.

Song, C., Woodcock, C., Seto, K., Lenney, M., Macomber, S., 2001. Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? *Remote Sensing of Environment* 75: 230-244.

Steuter, A., Grygiel, C., Biondini, M., 1990. A synthesis approach to research and management planning: The conceptual development and implementation. *Natural Areas Journal* 10: 61-68.

Torrence, C. and Compo, G., 1998. A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society* 79: 61-79.

Uuemaa, E., Antrop, M., Roosaare, J., Marja, R., Mander, U., 2009. Landscape metrics and indices: An overview of their use in landscape research. *Living Reviews in Landscape Research* 3: 5-28.

Vitt, P., Havens, K., Kendall, B., Knight, T., 2009. Effects of community-level grassland management on the non-target rare annual *Agalinis auriculata*. *Biological Conservation* 142: 798-805.

Wang, Q. and Malanson, G., 2007. Patterns of correlation among landscape metrics. *Physical Geography* 28: 170-182.

Zar, J., 1999. *Biostatistical Analysis 4th edition.* Prentice-Hall Inc., U.S.A.

Zhang, C. and Guo, X., 2007. Measuring biological heterogeneity in the northern mixed prairie: a remote sensing approach. *Canadian Geographer* 51: 462-474.

Zhang, C. and Guo, X., 2008. Monitoring northern mixed prairie health using broadband satellite imagery. *International Journal of Remote Sensing* 29: 2257-2271.

3. Relationship between the Spatial Configuration of Bare Sand Habitat and *Dalea villosa* Occurrence

3.1 Abstract

Identification of potentially suitable habitat is a key first step in critical habitat prediction for species protected by federal laws. Therefore, the objective of this research was to study the relationship between habitat configuration and hairy prairie-clover occurrence in order to predict suitable and unsuitable bare sand habitat across the study site. Bare sand patches were extracted from a land cover classification of the study site and several patch scaled metrics were calculated to characterize habitat spatial structure. Binary logistic regression was used to determine which metrics were significantly correlated with hairy prairie-clover occurrences. The logistic regression equation was subsequently used to predict suitable and unsuitable bare sand habitat for hairy prairie-clover based on the probability of occupancy. Results showed that about 29% of the variation in bare sand patch occupancy could be explained by the size, shape, and degree of isolation of a sand patch as well as the amount of vegetation on a sand patch in the early growing season. Based on these variables, 18.8% of bare sand patches in the study site were predicted to be unsuitable hairy prairie-clover habitat, 45.7% were predicted to be marginally unsuitable, 32.7% were predicted to be suitable, and 2.8% were predicted to be marginally suitable. Therefore, there is evidence that the spatial configuration of bare sand areas may not be the dominant factor limiting the abundance of hairy prairie-clover on the landscape.

3.2 Introduction and Background

Many methods for habitat suitability modelling now occur, each with differences in how they select variables, weight the contribution of variables, include interaction factors for variables, and predict the distribution of occurrences (Elith et al., 2006). One challenge with predictive modelling is that many species records include presence-only data with no reliable records of absences. Popular methods such as GARP, BioClim, and Maxent have proven accuracy for modelling when presence-only data is available for a species (Elith et al., 2006; Parolo et al., 2008). However, evaluating model performance for presence-only methods still requires the use of presence/absence data. Evaluating model performance with presence-only data limits the options for the data set and the power of statistical evaluations (Elith et al., 2006). Therefore,

where reliable presence/absence data sets are available, it is best to choose methods such as generalized linear models, generalized additive models, regression-based models, or resource selection functions (Hirzel and Le Lay, 2008; Parolo et al., 2008). Further, most theoretical models assume that response curves are either sigmoid or guassian, although in real ecological systems guassian curves may be rare due to factors such as thresholds along an environmental gradient or interspecific interactions. Therefore, regression-based habitat suitability models, such as logistic regressions, may be a better choice because they can fit curves ranging from parametric functions to less constrained shapes (Hirzel and Le Lay, 2008).

For plant species occupying discrete habitat niches, such as is the case with hairy prairie-clover, habitat suitability modeling can be effectively applied (Freckleton and Watkinson, 2002). The ecological niche concept is based on the theory that species only thrive in discrete ranges of environmental conditions (Hirzel and Le Lay, 2008). However, due to abiotic and biotic interactions with the surrounding environment a species may be constrained to occupy only a subset of its fundamental niche. This subset can be described as the realized niche. Habitat suitability modeling is an application of the ecological niche theory because it aims to predict the probability of occurrence of a species based on environmental conditions. The geographical distribution of a species generally follows three constraints: 1) the local environment allows the population to grow, 2) interactions with other local species allow the species to persist, and 3) the location is accessible given the dispersal abilities of the species (Hirzel and Le Lay, 2008). Therefore, habitat suitability modeling can only reconstruct the realized niche for a species based on the environmental variables present in the locations it occupies (Williams et al., 2009).

When dealing with plant species, it is important to consider that they are immobile, exhibit a strong spatial structure through specific habitat preferences, and their dispersal is restricted (Freckleton and Watkinson, 2002). Therefore, the geographical distribution of a plant species is heavily dependent on the geographical distribution of its realized niche. Thus, the above three constraints can be equated to functions of patch area, patch qualtiy, and patch isolation, respectively (Rizkalla et al., 2009). The three most basic processes determining the survival of a local population patch are dispersal, establishment, and persistence (Freckleton and Watkinson, 2002). It has been adequately documented that dispersal is mainly affected by patch isolation and matrix quality; establishment depends on patch size, number of patches, and patch quality; and persistence depends on patch size and patch quality (Kolb and Diekmann, 2005;

Kindlmann and Burel, 2008). Sensitivity to these spatial factors has been noted to be higher in rare species and habitat specialists as opposed to generalist species because they tend to have lower reproductive outputs, have shorter lived seeds, are smaller in size, and are geographically restricted (Kolb and Diekmann, 2005; Leach and Givnish, 1996).

Many habitat suitability models still fail to be successfully applied to real ecological problems because they fail to address spatial structure through assuming that all patches are equivalent in size and quality, and that all local populations are equally accessible by dispersers. Considering that all plant populations are spatially structured, habitat pattern must be incorporated into habitat suitability models to allow for an ecologically meaningful interpretation (Fahrig and Merriam, 1994). However, a distinction between habitat parameters that can be mapped and measured versus habitat parameters that are ecologically relevant must be made (Gustafson, 1998).

Understanding landscape pattern is important for understanding the effects of habitat change on rare flora populations such as hairy prairie-clover. Changes in landscape pattern can lead to changes in population pattern which can affect population processes and the long term survival of plant species (Kolb and Diekmann, 2005; Kindlmann and Burel, 2008). For hairy prairie-clover, the dominant habitat requirement appears to be an element of bare sand cover resulting in the plants confinement to geographical features such as parabolic sand dunes, stabilized blow-outs, dune depressions, and sand flats (Smith, 1998). It has been identified that factors such as climate change, vegetation encroachment, invasive species encroachment, anthropogenic sources and land use changes, and loss of natural disturbance by bison and fire have contributed to the reduction in sand dune activity and bare sand areas within the prairies and changes in the pattern of these landscape elements (Wolfe et al., 1995; Vance and Wolfe, 1996; Hugenholtz and Wolfe, 2005). The association of hairy prairie-clover with these types of bare sand areas makes it particularly vulnerable to changes in habitat pattern as incurred by the above factors. It has also been identified that vegetation and invasive species encroachment of bare sand and sand dune areas, loss of grazing and fire disturbance pressures, and changes in land use patterns are the main threats to the survival of hairy prairie-clover (Smith, 1998). Thus, it is hypothesised that the dominant factor limiting the abundance of hairy prairie-clover on the landscape is the spatial configuration of bare sand areas in context to other landscape elements.

3.2.1 Objective and Research Questions

The objective of this research is to study the relationship between habitat configuration and hairy prairie-clover occurrence. The purpose of this research is to analyse habitat pattern to gain an understanding of the spatial structure of bare sand habitat for hairy prairie-clover and to predict potentially suitable and unsuitable habitat for this plant. To achieve the second objective, the desired output is a habitat suitability map for hairy prairie-clover. In spatial habitat analysis, a distinction between habitat parameters that can be mapped and measured versus habitat parameters that are ecologically relevant is not always made (Gustafson, 1998). Therefore, objective two consisted of three specific research questions:

1) What is the spatial pattern of *intra-patch* characteristics for occupied and unoccupied bare sand habitat?

2) What is the spatial pattern of *inter-patch* characteristics for occupied and unoccupied bare sand habitat?

3) What spatial configuration of bare sand habitat best explains and predicts occupancy by hairy prairie-clover?

3.3 Methods

The second objective of this research can be achieved through analysis of patch scaled metrics for potential bare sand habitat. Correlation of bare sand patch characteristics with hairy prairie-clover occurrences can answer objective two (Figure 3.1).

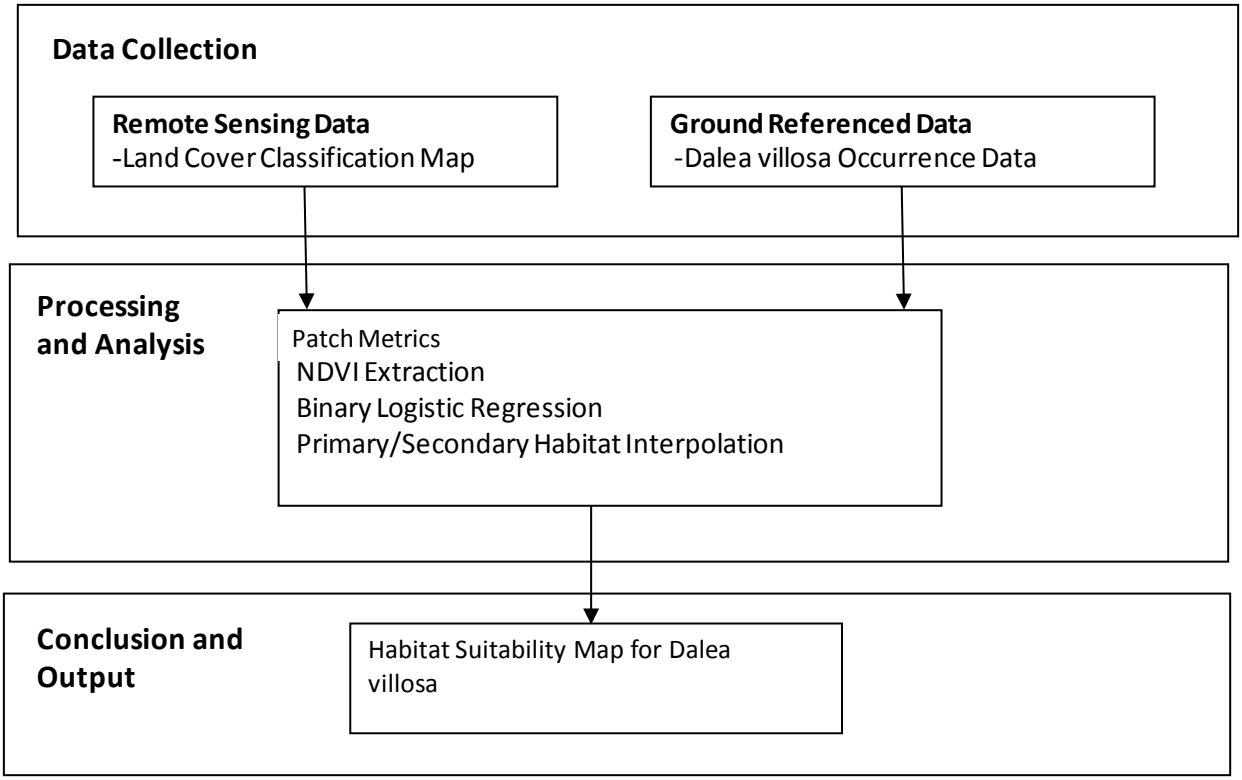


Figure 3.1: Flow chart of methods for objective two.

3.3.1 Data Set

3.3.1.1 Land Cover Classification

A SPOT5 panchromatic image (wavelengths 490-690 nm) at 2.5 m spatial resolution was acquired on April 13, 2007. This imagery was selected for its high spatial resolution. Three SPOT5 multispectral images (green, red, NIR, and MIR bands) at 10 m spatial resolution were acquired on May 27, 2009; July 12, 2009; and August 16, 2007. These images were selected for the spectral bands and temporal resolution.

A multi-temporal, multi-resolution land cover classification was carried out using object-oriented methods. Class boundary delineation was completed using a multiresolution segmentation algorithm at a scale parameter of ten, with shape and compactness weighted at 0.1 and 0.5 respectively. Following segmentation, the landscape was classified in to the five main land cover types present in the study site: aspen, shrub, juniper, grassland, and bare sand with a sixth class, other, added during post classification editing. A nearest neighbour classifier was set up using twenty six training data points collected in the field for each class and class membership values were assigned based on the mean reflectance within the panchromatic, red, NIR, and MIR

bands of an object. A hierarchical rule set was developed to maximize spectral and shape separability between classes.

The raster data format of the land cover classification was converted to vector data (polygon shape file), retaining the original boundaries of raster data, in ArcMap 9.2. Bare sand patches were extracted from the classification and dissolved to create a separate layer of the class bare sand. The remote sensing data set consisted of the bare sand layer.

3.3.1.2 Occupancy Data for Hairy Prairie-Clover

Occupancy data for hairy prairie-clover was collected from about 200 person-days of field work between 2007 and 2010. A previous coarse resolution land cover map was used to identify 351 potential habitat patches within which to restrict ground-based searches for hairy prairie-clover. Potential habitat patches included a 30 m buffer on grassland and shrub surrounding all classified bare sand patches (Environment Canada, unpublished data). Systematic searches of these habitat patches were conducted from July to September each year to allow for easy detection as the plant is in bloom during this time period. Systematic searches targeted bare sand areas and included a complete search of the bare sand patch and a buffer area of 30 m surrounding each patch. If plants were found, the outer boundary of the population was tracked using GPS (+/- 5 m error) to record the spatial extent of the patch. In spatial analyses, pseudo replication of a variable will depend on the distance between the sample points, such that a set of closely spaced sample points will be spatially autocorrelated and thus decrease the effective sample size (Parolo et al., 2008). Therefore, to minimize spatial autocorrelation issues, any hairy prairie-clover plants that were found within a 30 m proximity of one another were counted as the same patch/population.

This polygonal occupancy data was overlayed on the bare sand layer extracted from the current land cover classification (see section 3.3.1.1 above) to create an attribute table for bare sand patches where hairy prairie-clover was present or absent. Presence/absence data from 2007-2009 was used as training data in model building while presence/absence data collected in 2010 was kept separate for use as validation data in accuracy assessment (Figure 3.2). Hirzel and Guisan (2002) found that regular and equal-stratified sampling designs were the most accurate and robust for habitat suitability and presence/absence predictions. However, gradients that exert major control over a species distribution should be used to stratify sampling and increase

efficiency (Hirzel and Guisan, 2002). Therefore, in this case it was most appropriate to use systematic searches of targeted bare sand areas in data collection.

A. B.

Figure 3.2: Spatial distribution of known hairy prairie-clover presences and absences. The grey polygons represent known presences and the black polygons represent known absences. A. Presence/Absence data used in model building. B. Presence/Absence data used in model validation.

3.3.2 Metrics for Bare Sand Patch Configuration

To determine inter and intra patch spatial characteristics of bare sand patches, patch scaled metrics were calculated in ArcMap 9.2 for bare sand areas. Metrics were chosen based on the premise that the likelihood of a species occupying a particular habitat niche is a function of patch area and patch isolation (Rizkalla et al., 2009). In plant species, dispersal can be affected by patch isolation, establishment by the availability of suitable habitat and habitat quality, and persistence by patch size (Kolb and Diekmann, 2005). Therefore, six metrics were calculated for bare sand patches: area, shape, fractal dimension, Euclidean nearest neighbour distance, proximity count, and proximity percent. The normalized difference vegetation index (NDVI) was also calculated as a measure of habitat quality.

Patch area and perimeter were calculated in ArcMap 9.2 using Hawth's Table Tools. Shape was calculated as

$$SHAPE = \frac{0.25\, pij}{\sqrt{aij}} \tag{3.1}$$

where *pij* is the patch perimeter and *aij* is the patch area, adjusted by the constant 0.25 to adjust for the square standard in raster data (McGarigal and Marks, 1995). Shape will equal one if a

patch is perfectly square and increase without limit as patch shape becomes more irregular (McGarigal and Marks, 1995). It is a unitless measure. Fractal dimension, another measure of patch shape, was calculated as

$$FRACT = \frac{2\ln(0.25pij)}{\ln aij} \quad (3.2)$$

where *pij* is the patch perimeter, *aij* is the patch area, 0.25 is the constant for the square standard in raster data, and *ln* is the natural logarithm (McGarigal and Marks, 1995). Fractal dimension will vary between a value of one and two (unitless), where one signifies a square patch and two is the upper limit of shape irregularity (McGarigal and Marks, 1995).

The Euclidean nearest neighbour (ENN) distance was calculated as the straight line distance (meters) between the centroids of the closest neighbouring patches. Bare sand patch centroids were created in ArcMap 9.2 based on the gravimetric center of patches. Proximity measures, as a representation of patch isolation, were calculated in ArcMap 9.2 based on a 30 m buffer zone surrounding bare sand patches. Proximity count is the number of bare sand patches that fall within the 30 m buffer zone of the target sand patch. Proximity percent is the percentage of the area of the 30 m buffer zone that is of the bare sand class. These measures were modified from the proximity index by McGarigal and Marks (1995) to create a more meaningful output. It was pre-determined that any hairy prairie-clover plants that were found within a 30 m proximity of one another would be counted as the same population patch (see section 3.3.1.2 above). Given that hairy prairie-clover spreads mainly by seed (Godwin and Thorpe, 2007) it is plausible that dispersal distances fall within this proximity. Therefore, any bare sand patches within a 30 m proximity of one another were considered as the same habitat patch.

In plant species, one aspect of habitat quality is the competition for resources with other flora species, especially in inferior competitors because plants are immobile (Hirzel and Le Lay, 2008). Therefore, it is important to consider the amount of vegetation growing on bare sand patches as one measure of habitat quality for hairy prairie-clover. NDVI was used as a surrogate measure of the amount of vegetation on a bare sand patch. NDVI can be calculated for an image on a per-pixel bases as

$$NDVI = \frac{(NIR-Red)}{(NIR+Red)} \quad (3.3)$$

The difference in reflectance between the red band (absorption in vegetation, reflectance in soil) and NIR band (reflectance in vegetation and soil) will enhance vegetation cover across the

imagery. NDVI is one of the most popular vegetation indices because it can capture seasonal changes in vegetation growth, it is highly correlated with leaf area index, and it works best in low biomass conditions such as grassland, semi-arid, and arid regions (Jensen, 2005). NDVI was calculated for the SPOT5 multispectral image from August 2007 and the SPOT5 multispectral image from May 2009 using EASI Modelling in PCI Geomatica software. NDVI from August was chosen because during the late growing season, the maximum amount of vegetation coverage on a bare sand patch will show up (Jensen, 2005). NDVI from May was also chosen to show the amount of vegetation on a bare sand patch in the early growing season. Competition for resources in the early growing season can affect seedling establishment and persistence which is important for population growth (Hirzel and Le Lay, 2008; Schleuning and Matthies, 2008). The average NDVI for each sand patch was used.

3.3.3 Statistical Analysis and Model Building

Bare sand patches were coded based on hairy prairie-clover occupancy data where zero was unoccupied, one was occupied, and two was unknown. The data set consisted of 314 occupied and 314 unoccupied bare sand patches randomly selected from the 2007-2009 occurrence data. Using a binary response variable (presence/absence) in logistic regression is a popular approach to habitat suitability modelling (Hirzel and Guisan, 2002). However, sampling of the data should be done in a way to reduce model bias. Hirzel and Guisan (2002) found that regular and equal-stratified sampling designs were the most accurate and robust for habitat suitability and presence/absence predictions. Therefore, equal sample sizes were chosen to represent occupied and unoccupied sand patches. The independent variables tested were area, shape, fractal dimension, Euclidean nearest neighbour distance, proximity count, proximity percent, mean NDVI August, and mean NDVI May. For all tests, $\alpha=0.05$ was used. Assumptions of normality were tested using the one-sample kolmogorov-smirnov test. For independent variables in normal distribution, the independent-samples t-test was used to test for a significant difference between occupied and unoccupied bare sand patches. Assumptions of equal variances were tested using Levene's test for equality of variances. The independent-samples mann-whitney U test was used to test for a significant difference between occupied and unoccupied bare sand patches for variables not in normal distribution (Zar, 1999).

Binary logistic regression was used to determine which independent variables were significantly correlated with hairy prairie-clover occurrences, because the dependent variable

was binary (occupied or unoccupied), but the independent variables were continuous (Zar, 1999). Logistic regression is advantageous because it provides an equation for predicting the probability of occupancy on a bare sand patch, it uses multivariate comparisons between occupied and unoccupied sand patches, and it can identify the variables most associated with the occurrence of hairy prairie-clover (Wolken et al., 2001). The stepwise selection of independent variables can be useful for selecting an optimal model because the full model may be oversized, overfitted, or redundant (Parolo et al., 2008). Therefore, a forward stepwise likelihood ratio method was used (enter 0.05, removal 0.10). Correlation between independent variables can influence the final model because the selection and contribution of a vairiable depends on other variables already in the model (Wolken et al., 2001). Therefore, to avoid multicolinearity, only shape was used in the logistic regression and not fractal dimension. Mean NDVI August was also left out of the model because there was no significant difference between occupied and unoccupied bare sand patches for this variable ($p>0.05$). Variables considered in the model were area, shape, Euclidean nearest neighbour distance, proximity count, proximity percent, mean NDVI May, and all possible interactions. The hosmer-lemeshow test was used to test the assumption that there is a linear relationship between the independent variables and the log odds of the dependent variable (Zar, 1999).

3.3.4 Habitat Suitability Predictions and Model Validation

The logistic regression equation was subsequently used to calculate the probability of occupancy for all bare sand patches. Probability of occupancy was calculated as

$$P = \frac{Odds}{(1+Odds)} \tag{3.4}$$

where odds was defined as

$$Odds = e^{a+bx} \tag{3.5}$$

with $a+bx$ being the logistic regression equation. Hairy prairie-clover occupancy data from 2010 was used for accuracy assessment of model predictions. The data set consisted of 377 unoccupied and 131 occupied sand patches. Hairy prairie-clover occurrence polygons were overlain with the bare sand layer to create a contingency table of the model predictions (predicted suitable/unsuitable) against actual observations (observed presence/absence) (Table 3.1).

The Kmax index is the highest kappa value achieved when calculating the kappa coefficient at all threshold values between zero and one (Elith et al., 2006; Hirzel et al., 2006). It is a threshold-independent evaluator. Contingency tables were constructed for threshold values 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, and 0.9. Threshold values are defined in Table 3.2. The kappa coefficient was calculated for each threshold value as

$$KIA = \frac{N\sum xii - \sum xi.x.i}{N^2 - \sum xi.x.i} \tag{3.6}$$

where *x* and *N* are counts of evaluation points as defined in Table 3.1 (Hirzel et al., 2006). The sensitivity of prediction (percentage of presences correctly predicted) was calculated from the contingency tables as

$$sensitivity = \left({}^{xii}/_{x.i}\right) * 100 \tag{3.7}$$

where *xii* and *x.i* are defined in Table 3.1. The specificity of prediction (percentage of absences correctly predicted) was calculated from the contingency tables as

$$specificity = \left({}^{xss}/_{x.s}\right) * 100 \tag{3.8}$$

where *xss* and *x.s* are defined in Table 3.1. The overall percentage correctly predicted was calculated from the contingency tables as

$$Overall\ \% = \left[{}^{(xss + xii)}/_{N}\right] * 100 \tag{3.9}$$

where *N* is defined in Table 3.1.

Table 3.1: Contingency table for calculation of KIA, sensitivity, specificity, and overall % (Hirzel et al., 2006).

		Predicted		
		Unsuitable (0)	**Suitable (1)**	**sum**
Observed	**Absence (0)**	*xss*	*xis*	*x.s*
	Presence (1)	*xsi*	*xii*	*x.i*
sum		*xs.*	*xi.*	*N*

Table 3.2: Predicted suitable and unsuitable hairy prairie-clover habitat based on probability of occupancy for threshold values varying between zero and one. P represents the probability of occupancy for a bare sand patch and varies between zero and one.

Threshold Value	Suitable Habitat	Unsuitable Habitat
0.1	P>=0.1	P<0.1
0.2	P>=0.2	P<0.2
0.3	P>=0.3	P<0.3
0.4	P>=0.4	P<0.4
0.5	P>=0.5	P<0.5
0.6	P>=0.6	P<0.6
0.7	P>=0.7	P<0.7
0.8	P>=0.8	P<0.8
0.9	P>=0.9	P<0.9

The area under curve (AUC) is one of the most common threshold-independent methods for evaluating species' distribution models because it allows one to weight the risks of over and under predicting species' presences (Elith et al., 2006; Hirzel et al., 2006). For each threshold value 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, and 0.9 the true-positive ratio was plotted against the false-positive ratio and the area under the resulting curve was calculated. As well, a second curve was plotted with the true-negative ratio against the false-negative ratio for calculation of the area under the curve. The true-positive ratio was calculated from the contingency tables as

$$true\ positive = {}^{xii}/_{x.i} \tag{3.10}$$

as defined in Table 3.1 (Hirzel et al., 2006). The false-positive ratio was calculated from the contingency tables as

$$false\ positive = {}^{xis}/_{x.s} \tag{3.11}$$

as defined in Table 3.1 (Hirzel et al., 2006). The true-negative ratio was calculated from the contingency tables as

$$true\ negative = {}^{xss}/_{x.s} \tag{3.12}$$

as defined in Table 3.1 (Hirzel et al., 2006). The false-negative ratio was calculated from the contingency tables as

$$false\ negative = {}^{xsi}/_{x.i} \tag{3.13}$$

as defined in Table 3.1 (Hirzel et al., 2006). The AUC value will vary between zero (worse than random model) and one (best possible model), with values of 0.70 and greater indicating average to excellent model accuracy (Parolo et al., 2008).

The predicted-to-expected (P/E) ratio was the final evaluation method used to assess model accuracy because it is threshold-independent, it is an assessment of presence-only data, and it can be used to predict several levels of habitat suitability other than just suitable and unsuitable (Hirzel et al., 2006). Ten habitat suitability classes were defined based on the threshold values between zero and one (Table 3.3). For the P/E curve, ten classes with a class width of 0.1 is the optimum (Hirzel et al., 2006). Hairy prairie-clover presence data from 2010 was used in accuracy assessment. The data set consisted of 131 validation data points (presences only). For each habitat suitability class, the P/E ratio was calculated as

$$Fi = Pi/Ei \tag{3.14}$$

where *Pi* is the predicted frequency and *Ei* is the expected frequency (Hirzel et al., 2006). The predicted frequency was calculated as

$$Pi = \frac{pi}{\sum_{j=1}^{b} pj} \tag{3.15}$$

where *pi* is the number of validation data points that fall within the habitat suitability class *i* and $\sum pj$ is the total number of validation data points (Hirzel et al., 2006). The expected frequency was calculated as

$$Ei = \frac{ai}{\sum_{j=1}^{b} aj} \tag{3.16}$$

where *ai* is the area covered by class *i* and $\sum aj$ is the total area of the study site (Hirzel et al., 2006). In this case, the study site refers to bare sand patches only because only habitat is being considered and not total landscape. *Fi* was plotted against habitat suitability to produce the P/E curve. It is expected that a good model will show a linearly increasing curve with a low suitability class containing fewer presences than expected by chance ($Fi<1$) and high suitability classes having *Fi* increasingly higher than one (Hirzel et al., 2006). A threshold occurs where $Fi=1$ below which habitat is unsuitable and above which habitat is suitable (Hirzel et al., 2006). As well, additional thresholds can be placed at the steps in the curve to define several qualitative habitat suitability classes (Hirzel et al., 2006).

Table 3.3: Habitat suitability classes defined in the P/E curve. Ten classes with a class width of 0.1 is the optimum for calculating the P/E curve (Hirzel et al., 2006). P represents the probability of occupancy for a bare sand patch and varies between zero and one.

Habitat Suitability Class	Probability of Occupancy
0.0	p=0.0
0.1	0.0<P<=0.1
0.2	0.1<P<=0.2
0.3	0.2<p<=0.3
0.4	0.3<p<=0.4
0.5	0.4<p<=0.5
0.6	0.5<p<=0.6
0.7	0.6<p<=0.7
0.8	0.7<p<=0.8
0.9	0.8<p<=0.9
1.0	0.9<p<=1.0

Because the P/E curve did not match what was theoretically expected, k-fold cross-validation was used to assess any biases that may be caused by how the data was used. Hairy prairie-clover presence/absence data was randomly divided into three independent partitions (k=3) (Hirzel et al., 2006). K-fold cross-validation was applied using two of the partitions to calibrate the model and the left out partition in accuracy assessment, on only the retained variables from the original model (Hirzel et al., 2006). Therefore, two additional evaluations of model accuracy were produced to compliment the original accuracy assessment. Central tendency and variance of the P/E curve were evaluated from the mean, 95% confidence interval, and standard deviation to assess model robustness (Hirzel et al., 2006). The 95% confidence interval was calculated as

$$95\%\ CI = \bar{X} \pm t0.05(2), 2 * S\bar{X} \tag{3.17}$$

where $\bar{X}$ is the mean and $S\bar{X}$ is the standard error (Zar, 1999). Standard error was defined as

$$S\bar{X} = \frac{S}{\sqrt{n}} \tag{3.18}$$

where S is the standard deviation and n is the sample size (Zar, 1999). The standard deviation was calculated as

$$S = \sqrt{\frac{SS}{n-1}} \tag{3.19}$$

where the sum of squares (SS) was defined as

$$SS = \sum(Xi - \bar{X})^2 \tag{3.20}$$

The width of the confidence interval can assess the model's sensitivity to particular calibration points (Hirzel et al., 2006).

Bare sand patches were then re-coded as their predicted habitat suitability based on the probability of occupancy. The habitat suitability threshold for dividing suitable habitat (probability of occupancy above threshold) from unsuitable habitat (probability of occupancy below threshold) was chosen as the threshold value with the highest kappa coefficient, and where the true-positive ratio was maximized and the false-negative ratio was minimized (Hirzel et al., 2006). When choosing the habitat suitability threshold, the goal was to minimize the false-negative ratio (predicting that a sand patch is unoccupied when it actually is occupied) so as not to miss valuable hairy prairie-clover habitat. Additionally, four habitat suitability classes were defined from the steps in the P/E curve (Hirzel et al., 2006).

3.4 Results and Discussion

3.4.1 Inter-Patch and Intra-Patch Spatial Characteristics for Potential Bare Sand Habitat

Independent-samples t-tests and Mann-Whitney U tests revealed that there was a significant difference between occupied and unoccupied sand patches for most of the variables selected (Table 3.4). Occupied sand patches were significantly larger in size, more elongated shaped, less isolated, and less vegetated in the early growing season ($p<0.05$). There was no significant difference between the NDVI of occupied and unoccupied sand patches in the late growing season ($p>0.05$). Therefore, there is no evidence that late season vegetation cover on sand patches significantly affects the occurrence of hairy prairie-clover on bare sand areas. Although, there is evidence that vegetation cover on sand patches in the early growing season affects the occurrence of hairy prairie-clover. The mean NDVI in May was significantly lower on occupied sand patches as opposed to unoccupied sand patches (Table 3.4). Seedling establishment and persistence in the early growing season can be affected by competition for resources with other species, therefore it may be advantageous to hairy prairie-clover to establish on bare sand areas with reduced vegetation cover in the early growing season (Hirzel and Le Lay, 2008). As well, different types of vegetation green-up at different times during the growing season, therefore the NDVI in May could be detecting differences in the type of vegetation on sand patches rather than differences in the amount of vegetation. For example, it is commonly observed in the field that juniper inhabits partially to heavily stabilized bare sand areas. Considering juniper is an

evergreen shrub it will be captured in the early season NDVI as opposed to species that have not undergone green-up at this time. The difference in NDVI may be capturing the difference in juniper occurrence between sites.

Unoccupied bare sand patches tended to be smaller, more round shaped, more isolated, and more vegetated in the early growing season. Unoccupied sand patches had closer neighbouring sand patches than occupied bare sand areas but, were surrounded by less neighbouring patches. The ENN distance may be a poor measure of isolation because larger patches will have a greater distance between their patch centroids compared to smaller patches. Thus, the value may be an artefact of the measurement and not a true measure of isolation. This also suggests that it is not only the distance between habitat patches but the amount of surrounding habitat that is important to hairy prairie-clover occurrence. Dispersal is mainly affected by patch isolation and matrix quality with small patches being more affected by the quality of the matrix than larger patches (Lord and Norton, 1990; Kolb and Diekmann, 2005; Kindlmann and Burel, 2008). However, small patches that more densely populate the matrix can influence the nature of the intervening matrix to decrease patch isolation (Lord and Norton, 1990). Therefore, the amount of bare sand habitat within a 30 m proximity of the target patch may be a more ecologically relevant measure of patch isolation than just the ENN distance when considering factors affecting dispersal. Given that hairy prairie-clover spreads mainly by seed (Godwin and Thorpe, 2007) it is plausible that dispersal distances fall within this proximity. Therefore, any bare sand patches within a 30 m proximity of one another could be considered as the same habitat patch.

Unoccupied sand patches were, on average, smaller than occupied sand patches. However, the average area of an unoccupied sand patch was closer to the mean patch size of the bare sand class than occupied sand patches (Table 2.5 and 3.4). The average area of occupied sand patches shows greater dispersion from the mean patch size of the bare sand class, perhaps more influenced by a few larger sand patches. However, both the average area of unoccupied and occupied sand patches fall within one standard deviation of the mean patch size for the bare sand class (standard deviation=665) (Table 2.5). The mean patch size for the bare sand class is influenced by a large proportion of the sand patches in the study site being relatively smaller in size, however the large standard deviation indicates high variability in the size of bare sand patches.

Table 3.4: Significant differences between occupied and unoccupied sand patches for chosen patch metrics.

Variables	Mean (Occupied)	Mean (Unoccupied)	Significance
Area (m²)	724	188	P<0.05
Shape	1.69	1.35	P<0.05
Fractal Dimension	1.16	1.11	P<0.05
ENN Distance (m)	32	28	P<0.05
Proximity (count)	4	3	P<0.05
Proximity (%)	15.76	12.34	P<0.05
NDVI May	0.38	0.41	P<0.05
NDVI August	0.65	0.65	P>0.05

3.4.2 The Spatial Configuration of Hairy Prairie-Clover Habitat

Binary logistic regression revealed that the size and shape of a sand patch, as well as the degree of isolation, and amount of vegetation in the early growing season were significantly correlated with hairy prairie-clover occurrences (Equation 3.21). Almost 29% of the variation in bare sand patch occupancy could be explained by the logistic regression equation (r^2=0.285 and p<0.05).

$$\ln(ODDS) = 4.069 - 0.013Area + 0.036ProxPbyShape - 13.582NDVIMay + 0.038NDVIMaybyArea \quad (3.21)$$

The variable interaction of proximity percent with shape was the most useful for differentiating between occupied and unoccupied sand patches (Table 3.5). Similarly, other studies have found that the isolation of habitat patches is most important in explaining species richness and community composition because the rescue-effect allows populations to persist in small patches if they are well connected (Helm et al., 2006). Adding proximity percent, shape, area, and NDVIMay significantly improved the model's ability to predict suitable and unsuitable bare sand habitat for hairy prairie-clover (Table 3.5).

Table 3.5: Results of logistic regression analysis of variables characterizing occupied and unoccupied sand patches. *Significant at p<0.01

Variable Entered	Model Chi-Square	Improvement Chi-Square
ProxPbyShape	113.56*	23.82*
NDVIMay	129.49*	15.93*
NDVIMaybyArea	139.94*	10.45*
Area	152.26*	12.32*

Sand patch area and NDVIMay could be significant because larger sand patches tend to be less vegetated which decreases the competition for resources with other plants. Biotic interactions between plants act at short distances, but plants are immobile and have limited dispersal, therefore it is advantageous for inferior competitors to occupy competition-free locations (Hirzel and Le Lay 2008). Although some vegetation does take hold on the slipface of large sand dunes, these areas are generally kept clear of vegetation due to wind erosion and lack of soil nutrients and moisture favourable to the growth of most plant species (Hugenholtz and Wolfe, 2005), thus providing favourable habitat and competition-free locations for hairy prairie-clover.

Considering that herbivores are one of the major vectors of seed dispersal within the study site (Godwin and Thorpe, 2007), proximity percent and sand patch shape can help explain the spatial distribution of hairy prairie-clover. For example, ungulates can exhibit specific dispersal behaviour such as the orientation of movement along landscape elements and the ability to choose between patches, which will affect dispersal and exchange rates of seeds between patches (Heinz et al., 2006). Ungulates, such as white-tailed deer, are more likely to prefer linearly shaped patches because they provide more preferential edge habitat and corridors for movement (Rieucau et al., 2007; Volk et al., 2007). Thus, the shape of a sand patch could indirectly affect the spatial distribution of hairy prairie-clover through foraging behavior within the landscape. It was found that occupied sand patches were significantly more elongated than unoccupied sand patches (Table 3.4).

As well, the percentage of the immediate surrounding area (30 m buffer) of a sand patch could indirectly influence the spatial distribution of hairy prairie-clover through its influence on seed dispersal. Dispersal and exchange rates of seeds between patches will flow more readily in well connected patches based on the behavioural rule where individuals will stay at the first patch they reach as opposed to travelling extended distances between patches (Heinz et al., 2006). Occupied sand patches were also significantly less isolated than unoccupied sand patches (Table 3.4). Therefore, the spatial distribution of hairy prairie-clover can be explained by three constraints: 1) Patch Area (Area): the local environment allows the population to grow, 2) Patch Quality (NDVIMay): interactions with other local species allow the species to persist, and 3) Patch Isolation (ProxP and Shape): the location is accessible given the dispersal abilities of the species (Hirzel and Le Lay, 2008).

3.4.3 Habitat Suitability Predictions

Choosing the habitat suitability threshold where the maximum kappa coefficient is achieved will maximize model accuracy (Hirzel et al., 2006). The habitat suitability threshold for determining suitable habitat from unsuitable habitat was set at 0.40 because the kappa coefficient was maximized and the sensitivity of prediction was substantially higher than that achieved at a threshold of 0.50 (Table 3.6 and Figure 3.3). If the probability of a sand patch being occupied by hairy prairie-clover was equal to or greater than 40%, the sand patch was predicted to be suitable habitat. If the probability of occupancy was less than 40% then the sand patch was predicted to be unsuitable habitat. The goal was to minimize the false-negative ratio by maximizing the sensitivity of predictions. A desirable habitat suitability threshold occurred at 0.40 when considering the risks of over and under predicting suitable habitat (Figure 3.3). At a threshold value of 0.40, the true-positive ratio is maximized while the false-negative ratio is minimized. At this threshold, 35.5% of the sand patches in the study site were predicted to be suitable hairy prairie-clover habitat and 64.5% were predicted to be unsuitable habitat (Figure 3.4).

The results of the accuracy assessment for habitat suitability predictions show an average level of model accuracy (Table 3.6 and Figure 3.3). The highest kappa value obtained for the logistic regression model was 0.25, indicating a fair level of model agreement with ground referenced data (Landis and Koch, 1977). As well, the Area Under Curve (AUC) value at 0.70 indicates an average level of model accuracy (Parolo et al., 2008). Elith et al., 2006 found similar AUC values in their study comparing the accuracy of different species' distribution models. Across all species in all regions studied and considering only the best methods, they found that 14% of the species had AUC values around 0.70. At the chosen habitat suitability threshold, about 61% of the time the model will correctly predict suitable and unsuitable habitat; about 80% of the time, occurrences are correctly predicted; and about 54% of the time, absences are correctly predicted (Table 3.6).

Table 3.6: Accuracy assessment of habitat suitability predictions using the MaxKappa method. KIA is the kappa coefficient. Threshold values are based on the probability of occupancy as derived from Equation 3.21. Threshold values are defined in Table 3.2.

Threshold	KIA	Sensitivity of Prediction (% of occurrences correctly predicted)	Specificity of Prediction (% of non-occurrences correctly predicted)	Overall %
0.1	0.02	100	3.7	28.5
0.2	0.06	97.0	14.3	35.6
0.3	0.13	90.8	30.5	46.1
0.4	0.25	80.2	53.8	60.6
0.5	0.25	61.8	67.9	66.3
0.6	0.20	45.8	75.3	67.7
0.7	0.19	34.8	82.8	70.5
0.8	0.14	22.7	89.1	72.0
0.9	0.10	16.7	92.0	72.4

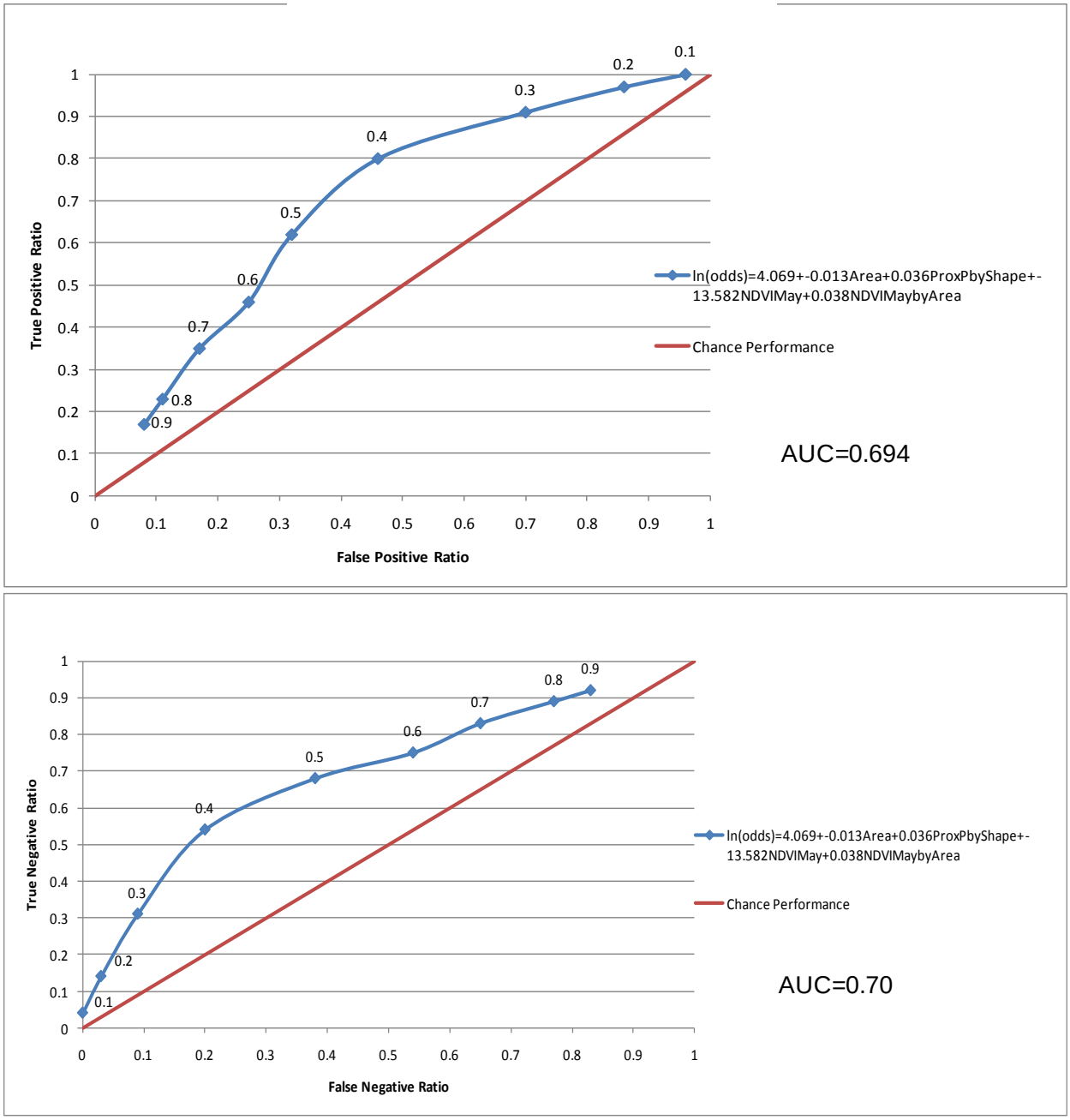


Figure 3.3: Accuracy assessment of habitat suitability predictions using the Area Under Curve (AUC) method. The red line represents a model that is no better than chance performance and the blue line represents the model derived from the logistic regression equation. The data points on the blue line represent the habitat suitability threshold values as defined in Table 3.2.

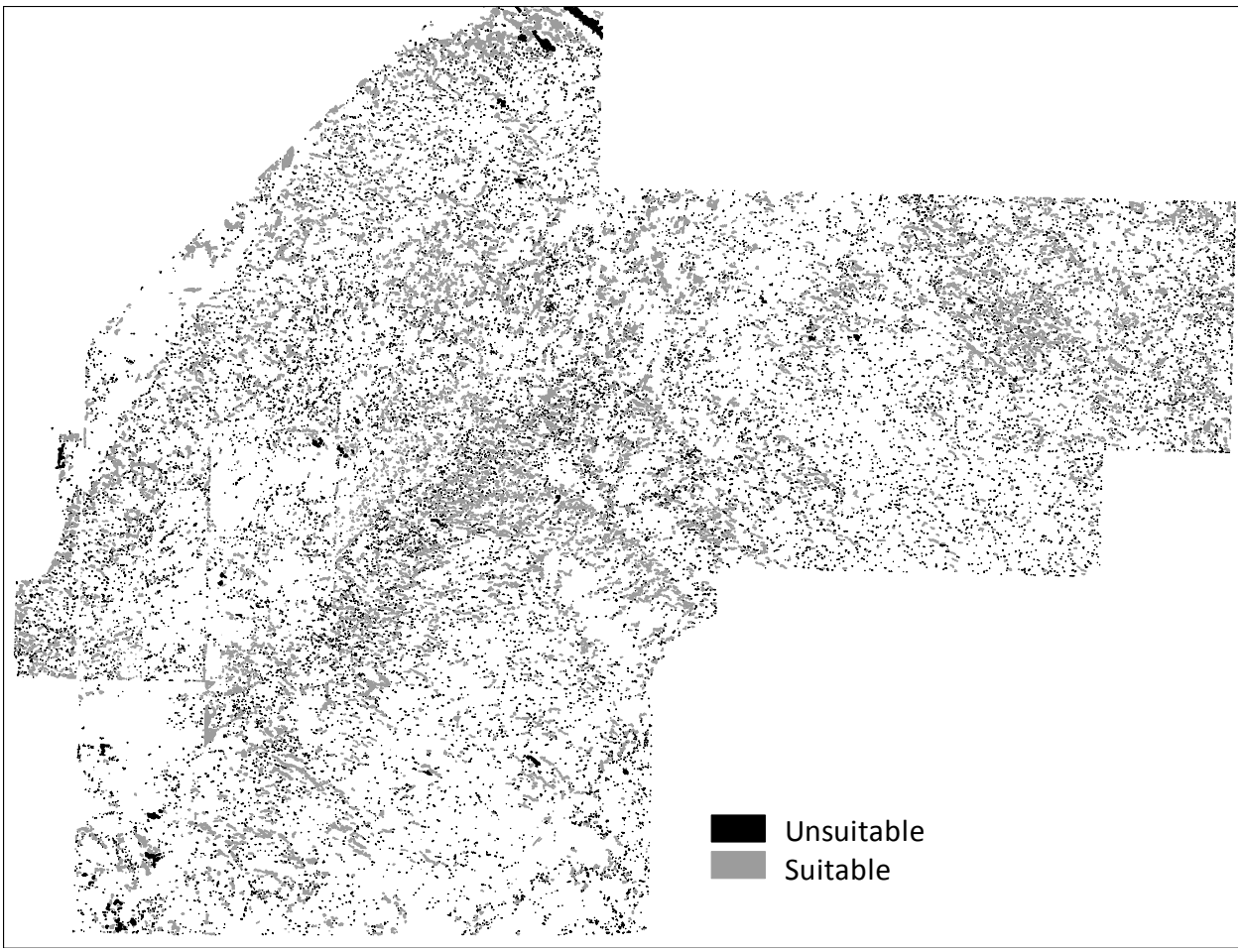


Figure 3.4: Habitat suitability predictions based on a threshold value of 0.40. Sand patches in grey were predicted to be suitable hairy prairie-clover habitat (probability of occupancy greater than or equal to 40%) and sand dunes in black were predicted to be unsuitable habitat (probability of occupancy less than 40%).

3.4.4 Relationship between Habitat Pattern and Population Pattern

The P/E curve was used to determine threshold values for four habitat suitability classes (Figure 3.5). The x-axis represents the probability of occupancy as defined in Table 3.3 and the y-axis represents the proportion of validation data that was predicted to fall into a certain habitat class to what would be expected to occur by chance (Figure 3.5). A low suitability class should contain fewer evaluation presences than expected by chance ($Fi<1$) and high suitability classes should have *Fi* increasingly higher than 1 (Hirzel et al., 2006). Similar to the results found with the MaxKappa and AUC methods, the habitat suitability threshold for separating suitable habitat from unsuitable habitat occured at 0.40 where the P/E curve crosses the $Fi=1$ line (Figure 3.5). Unsuitable habitat occurred across a probability of occupancy less than or equal to 20% where

there were fewer evaluation presences than expected by chance (Figure 3.5). Marginally unsuitable habitat occurred near the transition between suitable and unsuitable habitat at the *Fi*=1 line where the number of evaluation presences was close to what would be expected by chance (Figure 3.5). Suitable habitat occurred where the model predicted the probability of occupancy to be between 40% and 88%, because there were more evaluation presences than expected by chance (Figure 3.5). Where the model predicted the probability of occupancy to be greater than 88% (marginally suitable), validation data showed that occupancy of these types of sand patches actually dropped off which was not what was theoretically expected to occur (Figure 3.5).

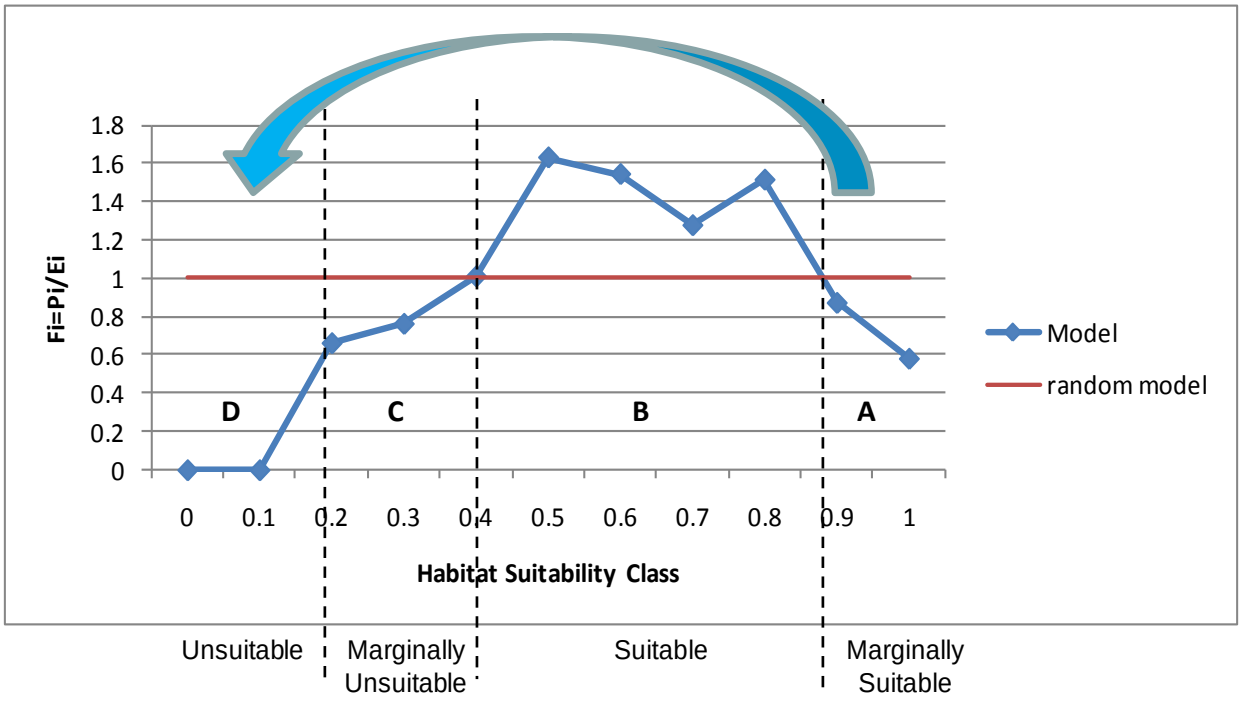


Figure 3.5: Four habitat suitability classes as derived from the P/E curve. The blue line represents the P/E curve derived from the logistic regression model and the red line represents a chance or random model. Fi is the predicted to expected ratio and the habitat suitability classes are defined in Table 3.3. The blue arrow represents the direction of natural succession when considering two variables from the logistic regression model: area and NDVIMay. A. active sand areas (Marginally Suitable habitat: increasing false absences). B. the transition state of a partially stabilized sand patch (Suitable habitat: mostly true presences). C. the remnant state of a heavily vegetated sand patch (Marginally Unsuitable habitat: increasing false presences). D. the altered state with the persistence of a new plant community (Unsuitable habitat: mostly true absences).

The shape of the P/E curve can be explained by how the logistic regression equation predicts the probability of occupancy. Looking at two variables, NDVIMay and area, the model predicts that as sand patches become larger and NDVIMay becomes lower the sand patch is

more likely to be occupied by hairy prairie-clover (Table 3.7). Following the shape of the P/E curve, this is true until a threshold is reached around a probability of 0.88 where actual occupancy drops back to being below predicted occupancy. Considering NDVIMay and area, natural succession on a sand patch would thus cause the probability of occupancy to decrease as colonizing species stabilize the bare sand area and gradually increase NDVIMay and decrease area (Figure 3.5). This is apparent when looking at the average area and average NDVIMay of each habitat suitability class (Table 3.7).

Landscapes are dynamic rather than static and are in a constant flux of change. The temporal scale of landscape change is different from the temporal scale of flora species' response to landscape change such that present species occurrences can still reflect past landscape patterns (Cousins et al., 2007; Helm et al., 2006). For example, Helm et al. (2006) found that current species richness in fragmented alvar grasslands was more significantly related to past areas and connectivities of habitat patches than current habitat pattern. As well, Cousins et al. (2007) found a significant relationship between present habitat area and current species richness, and habitat area 50 years ago and current species richness in grasslands in south-eastern Sweden. Eriksson (1996) found that remnant patches could exist within a population because there was a considerable time span between migration of new individuals to a patch and the final extinction of species in a patch during succession. Due to the delayed response of plants to habitat change, habitat patches can harbour species that prefer an environment that no longer exists (Cousins et al., 2007). Therefore, it is not uncommon that hairy prairie-clover was found to occupy marginally suitable habitat and marginally unsuitable habitat in this study.

Table 3.7: The average area and NDVIMay of four different habitat suitability classes. As sand patch size increases and NDVIMay decreases, the probability of occupancy increases.

Predicted Habitat Suitability	Probability of Occupancy	Mean NDVIMay	Mean Area (m^2)	% of Landscape
Unsuitable	p<=0.2	0.49	79.18	18.8
Marginally Unsuitable	0.2<p<=0.4	0.41	79.63	45.7
Suitable	0.4<p<=0.88	0.38	263.85	32.7
Marginally Suitable	p>0.88	0.41	2041.31	2.8

Considering the variables present in the logistic regression equation (Table 3.7) and the shape of the P/E curve, the four habitat suitability classes were derived based on the direction of

natural succession (Figure 3.5). The sand dunes found in the Dundurn area comprise four physiographic categories: active complexes, stabilized blow-outs, stabilized dunes, and dune depressions. When considering three dominant environmental gradients, soil-water holding capacity, edaphic characteristics, and degree of stabilization; stabilized blow-outs and dune depressions have closely related environmental niches while stabilized dunes and active complexes are distinct and follow a distinct environmental gradient between the two. Natural succession reflects the stabilization of active complexes, increasing the water retaining capacity and organic content of sandy soils, towards stabilized dunes. Comparatively, blow-outs represent a more advanced stage of succession due to decreased wind erosion allowing invasion by woody species such as creeping juniper. As well, the more mesic conditions in dune depressions as compared to the exposed slopes of stabilized dunes, allows late successional species to take over and develop a more complex vegetative community (Hulett et al., 1966).

On average, larger sized sand dunes are kept relatively clear of vegetation on their most active surfaces (the dune head, crest, and slip face) due to wind erosion and deposition of sand, and a lack of soil moisture and nutrients preventing growing conditions that are favourable to most species (Hugenholtz and Wolfe, 2005). Active complexes in the Dundurn sand-hills were found to have lower soil organic content, higher percentages of sand, and lower soil moisture (Hulett et al., 1966). Although these types of sand patches were predicted to have a probability of occupancy greater than or equal to 88% for hairy prairie-clover, accuracy assessment revealed that *Fi* was less than one (Figure 3.5 A). Therefore, these types of sand patches are only marginally suitable and the occurrence of false absences is increasingly higher in these types of sandy areas. False absences occur in these sandy areas because hairy prairie-clover tends to be absent from these potentially suitable sites (Figure 3.5 A). While currently they are not preferred habitat, hairy prairie-clover plants are able to colonize them allowing succession to progress towards an environment that is more favourable to the plant.

Hairy prairie-clover prefers to occupy the niche in the successional stage where bare sand areas are partially stabilized (Figure 3.5 B and Table 3.7). Thus, following the environmental gradient along the succession of active complexes to stabilized dunes (Hulett et al., 1966), hairy prairie-clover will be more commonly associated with stabilized dunes than active complexes (Figure 3.5). Hairy prairie-clover has been found to be positively associated with early invading species of open sand and other early successional species within the Dundurn sand-hills (Godwin

and Thorpe, 2007). Sparsely vegetated and frequently disturbed microsites such as partially stabilized sandy areas, favour plants with nitrogen fixing symbiosis because nitrogen fixation provides a net advantage on open, sunny, nitrogen poor sites (Leach and Givnish, 1996). Hairy prairie-clover occurrence in sandy areas provides a valuable link in the natural succession of sand dunes and bare sand patches because it is part of the legume family of which most legumes are nitrogen fixers (Hulett et al., 1966).

Hairy prairie-clover has also been found to exist in marginally unsuitable habitat (Figure 3.5 C) because nitrogen fixation provides soil nutrients which would otherwise be sparse in a bare sand patch allowing other species to colonize the area. For example, creeping juniper reaches its maximum abundance and is most associated with stabilized blow-outs in the Dundurn sand-hills. The main difference between stabilized dunes and stabilized blow-outs is the higher abundance of creeping juniper in the latter. As well, the most dominant species in dune depressions are *Stipa comata* and *Calamovilfa longifolia* which are two species that occur in the later stages of succession (Hulett et al., 1966). Hairy prairie-clover has been found growing in conjunction with creeping juniper on partially to heavily stabilized sandy areas, even though these sites are not necessarily favourable to the long term persistence of the plant (Godwin and Thorpe, 2007). The biotic interactions that occur with the co-existence of species, especially the existence of a superior competitor (such as the implied relationship between creeping juniper and hairy prairie-clover), will affect the fitness and behaviour of species such that the population will persist as a remnant patch with an extinction debt (Hirzel and Le Lay, 2008; Godwin and Thorpe, 2007; Helm et al., 2006). In marginally unsuitable habitats, such as blow-outs and depressions, an increase in false presences is observed. While hairy prairie-clover can occupy marginal habitats, these patches are not necessarily favourable for supporting the long term persistence of the population because natural succession proceeds towards a decrease in habitat suitability for hairy prairie-clover.

Model robustness can be assessed from the variance among the cross-validation curves across the entire range of habitat suitability to determine which parts of the model are most accurate. The standard deviation between the three models from K-fold cross-validation (Figure 3.6) can help determine which habitat suitability classes were predicted the best and where the model needs improvement (Hirzel et al., 2006). Unsuitable habitat was predicted with relatively high accuracy (Figure 3.6). The narrowness of the standard deviation shows that all three models

in K-fold cross-validation were in agreement for predicted unsuitable hairy prairie-clover habitat. As well, suitable habitat was predicted with high confidence as seen in the low variance between the three models from K-fold cross-validation (Figure 3.6). High variability between model predictions occurred for both marginal habitat types, indicating the models weakness in predicting marginally unsuitable and marginally suitable habitat for hairy prairie-clover. The greatest variation between models occurred at habitat suitabilities of 0.40 and 0.80, at the transition between suitable and unsuitable habitat where the curve crosses the *Fi*=1 line (Figure 3.6). This reflects the model's inability to account for false presences and false absences.

Considering that the spatial configuration of habitat could explain about 29% of the variation in patch occupancy, it is not surprising that the model was unable to account for false presences and false absences. Including other ecological factors which may influence the occurrence of false presences and false absences could help improve the predictive capability of the model. For example, a study by Parolo et al. (2008) which included ecological variables such as a topographic wetness index and direct incident solar radiation had an AUC value that was 0.18 higher than that obtained in this study. Variables such as elevation, slope angle, slope aspect, amount of incoming solar radiation, surface soil moisture, and surface wind speed have been identified as key ecological factors affecting population viability in plant species (Wolken et al., 2001; Parolo et al., 2008). Adding these ecological variables to the model could further explain hairy prairie-clover presence/absence on sand patches.

False presences and false absences occur in populations because species can occupy unsuitable habitat and/or be absent from suitable habitat. This is due to several reasons, namely 1) the difference in temporal scale between landscape change and flora species' response to change, and 2) species occupancy can only reflect the realized niche and not the fundamental niche due to abiotic and biotic environmental factors and interactions (Hirzel and Le Lay, 2008; Williams et al., 2009; Cousins et al., 2007). Therefore, the transition between suitable and unsuitable habitat cannot be modelled with high certainty. Further, while statistical methods can determine statistical significance between the mean of a patch metric, ecological significance is harder to detect because many ecological relationships are threshold dependent where only changes in pattern at the threshold are ecologically significant (Gustafson, 1998). For example, the extinction of long-lived plants would show an S-shaped curve representing thresholds where the effect on a species suddenly increases. Extinction would progress slowly at the beginning and

increase dramatically once a time or area threshold had been reached before levelling out after the extinction debt had been equalized (Cousins et al., 2007; Helm et al., 2006). Therefore, hairy prairie-clover can still be found in marginal habitats where successional change has not yet reached a threshold capable of eliciting a response in the plant's spatial occurrence (Figure 3.5). In marginally unsuitable habitat, this can result in remnant patches that exhibit an extinction debt (the proportion of the population that would go extinct if the distribution of the population caught up to match current habitat conditions) (Helm et al., 2006).

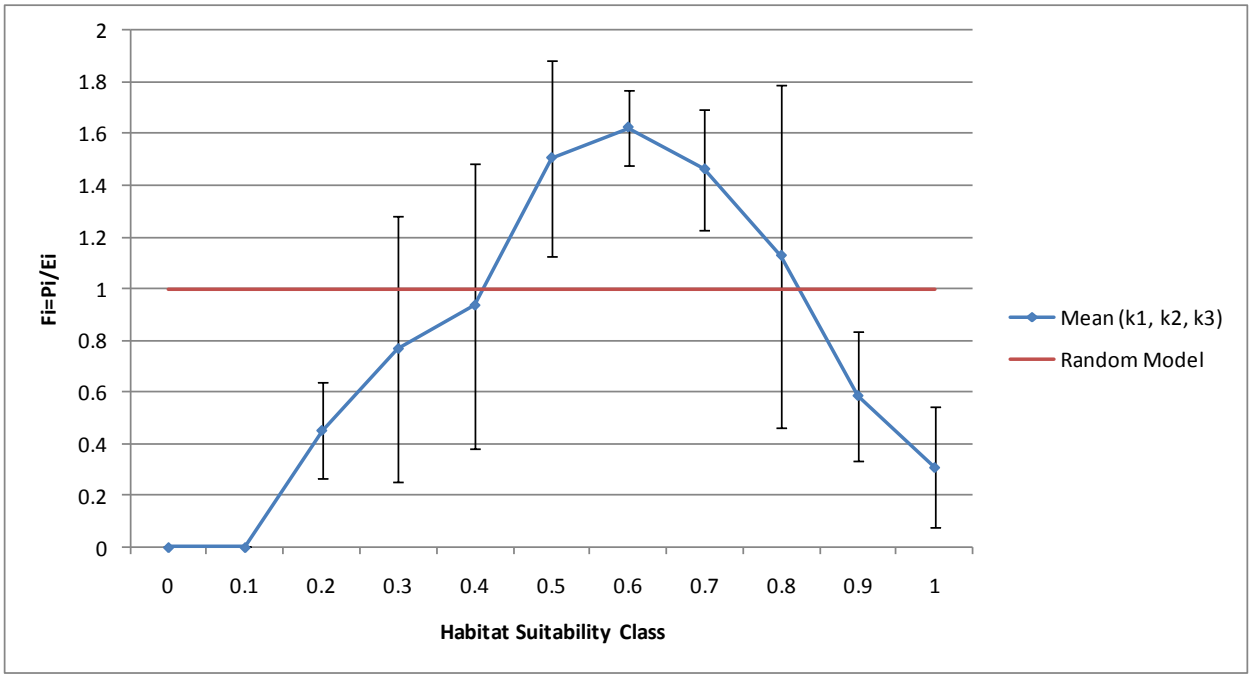


Figure 3.6: The standard deviation between the three models (K1, K2, K3) from k-fold cross-validation.

In the final habitat suitability map for hairy prairie-clover (Figure 3.7) almost 46% of the bare sand patches were predicted to be marginally unsuitable hairy prairie-clover habitat, and about 19% of the sand patches were predicted to be completely unsuitable (Table 3.7). About 33% of the sand patches were predicted to be suitable hairy prairie-clover habitat and only 2.8% were predicted to be marginally suitable (Table 3.7). Considering the state and transition model of succession in Figure 3.5, the majority of the bare sand patches in the study site exist as remnant patches undergoing succession. Therefore, there is evidence that in the absence of natural disturbance, the long term trend will be towards a decrease in habitat suitability for hairy prairie-clover.

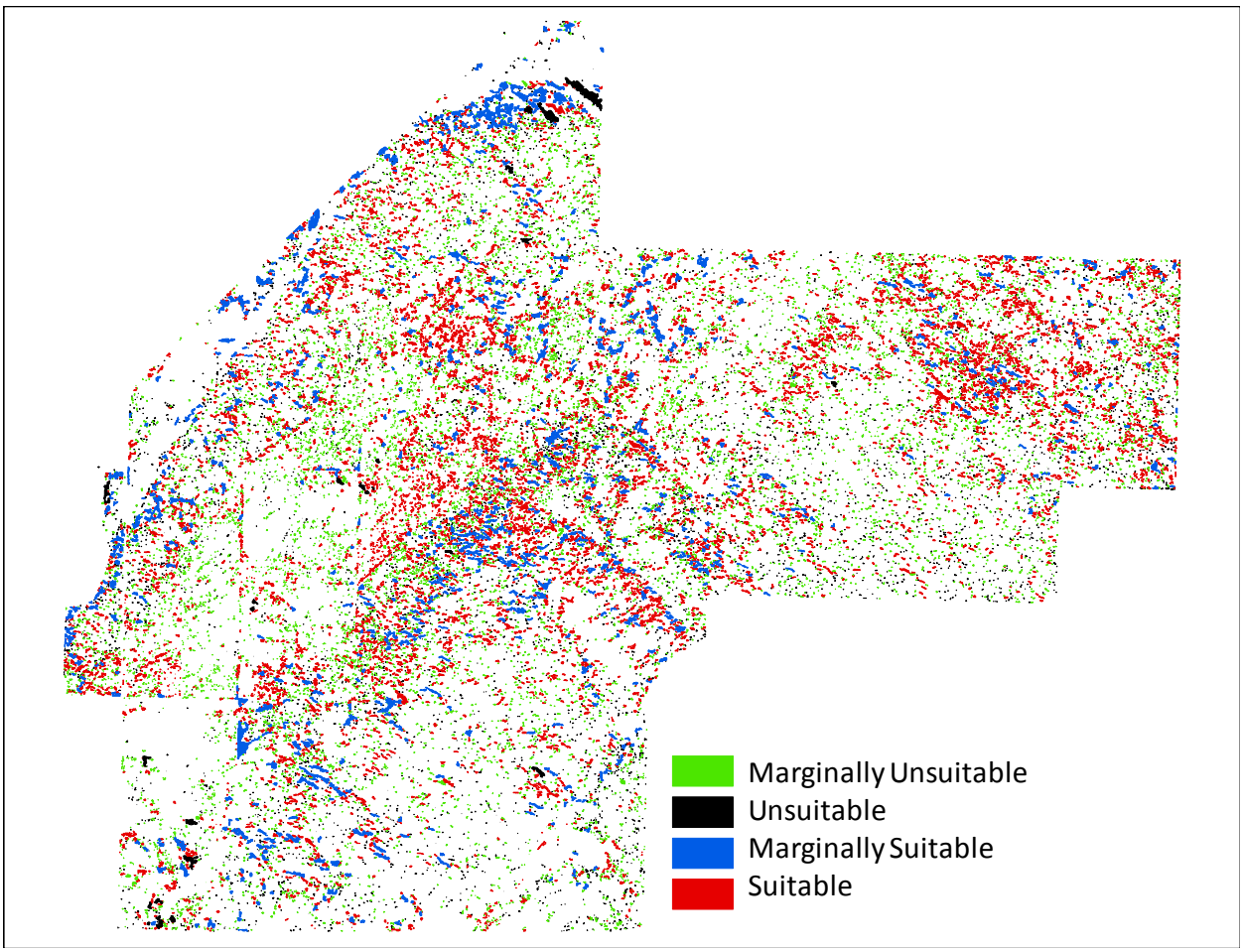


Figure 3.7: Habitat suitability map for hairy prairie-clover of bare sand patches. The white background consists of classes grassland, shrub, aspen, juniper, and other.

3.5 Conclusion

It has been demonstrated that habitat suitability for hairy prairie-clover can be predicted with an average level of model accuracy (AUC=0.70 and MaxKappa=0.25) when only considering factors of habitat spatial pattern. Four significant components of habitat pattern (area, shape, proximity percent, NDVIMay) were able to explain almost 29% of the variation in bare sand patch occupancy. Therefore, there is evidence that the spatial configuration of bare sand areas may not be the dominant factor limiting the abundance of hairy prairie-clover on the landscape. Landscape and habitat pattern may not be exerting as much control over population pattern as originally thought.

Contrary to what was theoretically expected by the logistic regression model alone, sand patches with a probability of occupancy between 40% and 88% were the most preferred habitat

by hairy prairie-clover when cross-validated with additional occurrence data. It was expected that the model would show a linearly increasing curve as the probability of occupancy increased (Hirzel et al., 2006). Deviation from the theoretical linear curve seen at probabilities greater than about 88% (Figure 3.5) could be a reflection of the time lag between the establishment of a species in an area (marginally suitable habitat with increasing false absences) and the persistence of that species as succession progresses (suitable habitat with true presenes). The majority of sand patches in the study site were predicted to be marginally unsuitable and unsuitable hairy prairie-clover habitat. There is evidence that successional trends in the study site are towards a decrease in habitat suitability for hairy prairie-clover. While it appears that current landscape and habitat structure is capable of supporting a large hairy prairie-clover population, the long term persistence of the species' may be threatened due to progressive formation of remnant patches and possible extinction debts (Helm et al., 2006). In the absence of natural disturbance to help maintain successional heterogeneity among sand patches, long term trends will be towards complete stabilization (unsuitable hairy prairie-clover habitat). Therefore, it is important to not only consider the structural component of habitat but the functional and compositional components as well.

3.6 References

Cousins, S., Ohlson, H., Eriksson, O., 2007. Effects of historical and present fragmentation on plant species diversity in semi-natural grasslands in Swedish rural landscapes. *Landscape Ecology* 22: 723-730.

Elith, J., Graham, C., Anderson, R., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R., Huettmann, F., Leathwick, J., Lehmann, A., Li, J., Lohmann, L., Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J., Peterson, A., Phillips, S., Richardson, K., Scachetti-Pereira, R., Schapire, R., Soberon, J., Williams, S., Wisz, M., Zimmermann, N., 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29: 129-151.

Eriksson, O., 1996. Regional dynamics of plants: a review of evidence for remnant, source-sink and metapopulations. *Oikos* 77: 248-258.

Fahrig, L. and Merriam, G., 1994. Conservation of fragmented populations. *Conservation Biology* 8: 50-59.

Freckleton, R. and Watkinson, A., 2002. Large-scale spatial dynamics of plants: metapopulations, regional ensembles and patchy populations. *Journal of Ecology* 90: 419-434.

Godwin, B. and Thorpe, J., 2007. Targeted surveys for plant species at risk in Elbow, Dundurn and Rudy-Rosedale AAFC-PFRA Pastures, 2006. Agriculture and Agri-Food Canada-Prairie Farm Rehabilitation Administration. *Saskatchewan Research Council Pub. No. 11997-1E07.*

Gustafson, E., 1998. Quantifying landscape spatial pattern: What is the state of the art? *Ecosystems* 1: 143-156.

Heinz, S., Wissel, C., Frank, K., 2006. The viability of metapopulations: individual dispersal behaviour matters. *Landscape Ecology* 21: 77-89.

Helm, A., Hanski, I., Partel, M., 2006. Slow response of plant species richness to habitat loss and fragmentation. *Ecology Letters* 9: 72-77.

Hirzel, A. and Guisan, A., 2002. Which is the optimal sampling strategy for habitat suitability modelling. *Ecological Modelling* 157: 331-341.

Hirzel , A., Le Lay, G., Helfer, V., Randin, C., Guisan, A., 2006. Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling* 199: 142-152.

Hirzel, A. and Le Lay, G., 2008. Habitat suitability modelling and niche theory. *Journal of Applied Ecology* 45: 1372-1381.

Hugenholtz, C. and Wolfe, S., 2005. Recent stabilization of active sand dunes on the Canadian prairies and relation to recent climate variations. *Geomorphology* 68: 131-147.

Hulett, G., Coupland, R., Dix, R., 1966. The vegetation of dune sand areas within the grassland region of Saskatchewan. *Canadian Journal of Botany* 44: 1307-1331.

Jensen, J., 2005. *Introductory Digital Image Processing 3rd Edition.* Prentice-Hall Inc., U.S.A.

Kindlmann, P. and Burel, F., 2008. Connectivity measures: a review. *Landscape Ecology* 23: 879-890.

Kolb, A. and Diekmann, M., 2005. Effects of life-history traits on responses of plant species to forest fragmentation. *Conservation Biology* 19: 929-938.

Koper, N., Mozel, K., Henderson, D., 2010. Recent declines in northern tall-grass prairies and effects of patch structure on community persistence. *Biological Conservation* 143: 220-229.

Landis, B. and Koch, G., 1977. The measurement of observer agreement for categorical data. *Biometrics* 33: 159-174.

Leach, M. and Givnish, T., 1996. Ecological determinants of species loss in remnant prairies.

Science 273: 1555-1558.
Lord, J. and Norton, D., 1990. Scale and the spatial concept of fragmentation. *Conservation Biology* 4: 197-202.
McGarigal, K. and Marks, B., 1995. FRAGSTATS: spatial pattern analysis program for quantifying landscape structure. Gen. Tech. Rep. PNW-GTR-351. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 122 p.
Parolo, G., Rossi, G., Ferrarini, A., 2008. Toward improved species niche modelling: Arnica Montana in the Alps as a case study. *Journal of Applied Ecology* 45: 1410-1418.
Rieucau, G., Vickery, W., Doucet, G., Laquerre, B., 2007. An innovative use of white-tailed deer (Odocoileus virginianus) foraging behaviour in impact studies. *Canadian Journal of Zoology* 85: 839-846.
Rizkalla, C., Moore, J., Swihart, R., 2009. Modelling patch occupancy: relative performance of ecologically scaled landscape indices. *Landscape Ecology* 24: 77-88.
Schleuning, M. and Matthies, D., 2008. Habitat change and plant demography: Assessing the extinction risk of a formerly common grassland perennial. *Conservation Biology* 23: 174-183.
Smith, B., 1998. COSEWIC status report on the hairy prairie-clover in Canada. *Environment Canada.*
Vance, R. and Wolfe S., 1996. Geological indicators of water resources in semi-arid environments: southwestern interior of Canada. In: Geoindicators: Assessing rapid environmental changes in earth systems. Berger, A.R. and Iams, W.J. (Eds.). A.A. Balkema, p. 251-263.
Volk, M., Kaufman, D., Kaufman, G., 2007. Diurnal activity and habitat associations of white tailed deer in tallgrass prairie of eastern Kansas. *Transactions of the Kansas Academy of Science* 110: 145-154.
Williams, J., Seo, C., Thorne, J., Nelson, J., Erwin, S., O'Brien, J., Schwartz, M., 2009. Using species distribution models to predict new occurrences for rare plants. *Diversity and Distributions* 15: 565-576.
Wolfe, S., Huntley, D., Ollerhead, J., 1995. Recent and late Holocene sand dune activity in southwestern Saskatchewan. *Geological Survey of Canada* 131-140.
Wolken, P., Sieg, C., Williams, S., 2001. Quantifying suitable habitat of the threatened western prairie fringed orchid. *Journal of Range Management* 54: 611-616.
Zar, J., 1999. *Biostatistical Analysis 4th edition.* Prentice-Hall Inc., U.S.A.

4. Integration and Synthesis

This chapter will synthesise the main findings and contributions of this research and relate it back to the broader context of the literature introduced in chapter one. Research limitations as well as future directions will also be discussed.

4.1 Significance and Contributions

Addressing the first research objective, it was concluded that the optimal spatial resolution for sand patch delineation and classification in the study site was between 2-5 m. The optimal temporal resolution for land cover classification included imagery from the early and late growing seasons to account for differences in phenology between land cover classes. In spatial habitat analysis, a land cover classification at an appropriate spatial and temporal resolution is necessary because landscape pattern and population pattern cannot have an ecologically meaningful relationship if they are not measured at the same scale (Gustafson, 1998). It has been adequately documented that the spatial variation observed in landscapes is a function of the variability in the scale at which abiotic and biotic factors operate such that a landscape can have more than one significant scale (He et al., 2007). As a result, the same landscape will be classified differently as data set resolution changes, such that when the purpose is to create a species specific habitat map, the resolution of the land cover classification must match the spatial scale at which the species operates in the environment. This study demonstrates the utility of conducting a land cover classification at a habitat specific spatial and temporal scale, which has rarely been done before. This allows for the integration of landscape spatial structure into habitat suitability models by considering that plant populations will be spatially structured from habitat pattern, to improve the application of metapopulation concepts to ecological problems (Fahrig and Merriam, 1994).

The second objective of this research expanded on this by using the land cover classification to correlate several metrics of habitat pattern with hairy prairie-clover presence/absence for use in habitat suitability modelling. Despite the theoretically overriding influence that landscape and habitat pattern have on population pattern, especially in plant species (Noss, 1990; Feckleton and Watkinson, 2002), most studies fail to reliably incorporate aspects of spatial structure into their habitat suitability models, until now. Results showed that almost 29% of the variation in sand patch occupancy could be explained by the size, shape, and

degree of isolation of a sand patch as well as the amount of vegetation on a sand patch in the early growing season. Occupancy tended to follow successional stages resulting in a non-linear curve. However, habitat suitability for hairy prairie-clover was still predicted with a relatively high level of model accuracy (AUC=0.70) when only considering factors of habitat spatial pattern. For example, Elith et al. (2006) found similar AUC values in their study comparing the accuracy of different species' distribution models. Across all species in all regions studied and considering only the best methods, they found that 14% of the species had AUC values around 0.70. Considering that all plant populations are spatially structured because they are immobile, have limited dispersal, and have specific habitat preferences (Freckleton and Watkinson, 2002); a habitat suitability map for hairy prairie-clover based on habitat spatial pattern provides a valuable tool for further land use management decision making in the study site.

The influence of landscape and habitat pattern on the spatial distribution of hairy prairie-clover has not been closely studied until now. There is evidence that the spatial configuration of bare sand habitat may not be the dominant factor limiting the abundance of hairy prairie-clover on the landscape. Other explanatory variables such as slope angle, slope aspect, amount of incoming solar radiation, surface soil moisture, nutrient content of soil, and community composition have been identified as key ecological factors affecting population viability in plant species (Wolken et al., 2001; Parolo et al., 2008). These factors have also been identified as possible explanations for the distribution of hairy prairie-clover at finer scales than this study (Godwin and Thorpe, 2007). However, almost 29% of the variation in patch occupancy could be explained by habitat spatial pattern and significant differences were found between occupied and unoccupied sand patches, giving evidence that habitat pattern is at least partially influential on the presence/absence of hairy prairie-clover at sites.

The demonstrated link between habitat pattern and population pattern in this study has important management implications when considering Noss's (1990) hierarchical perspective on ecosystem studies. For example, management of the structure, function, and composition of a landscape at the landscape or habitat scale has been demonstrated to be sufficient for the maintenance of rare and threatened species (Walker, 1995; Williams et al., 2006; Vitt et al., 2009). As well, mapping the spatial pattern of sand patches can be used for change detection and further analysis of sand dune stabilization which can be a valuable indicator of climate effects on the prairies (Hugenholtz and Wolfe, 2005). Montoring changes in bare sand habitat over time can

also provide valuable insight into how a plant responds to habitat loss and fragmentation, to further aid in land use management decisions. The methods used in this research are not just significant for habitat suitability modelling of hairy prairie-clover, but could also be applied to other habitat specialist plant species.

4.2 Limitations

One limitation of the land cover classification and spectral analysis is the difference in acquisition dates between ground referenced field data and remote sensing imagery. Due to limitations of budget and availability of quality remote sensing data (cloud cover less than 10%), existing images from 2007 were used along with newly acquired images from 2009 when field work was conducted. Although June, July, and August 2009 field data was used to represent the early, middle, and late growing season respectively, these phenological stages do not coincide with the same months every year. Therefore, while spectral analysis of field data gives a good estimation of the spectral response of land cover types in the different growing seasons, there are limitations when extrapolating this information to specific months within different years. Classification accuracy can also be affected by several problems inherent in grassland remote sensing: spatial variation in coverage will occur within a certain grassland type; spatial variation in the types of species present within a certain grassland type will occur; the existence of standing dead materials, litter, biological crust, and bare soil will affect the spectral signature of land cover classes; and identical vegetation types can vary in reflectance values due to differences in coverage or vigour (Herrera et al., 2009; Zhang and Guo, 2007; Geerken et al., 2005).

While the investigation of the spatial relationship between bare sand habitat for hairy prairie-clover and other land cover classes gave a good idea of the current landscape structure of the study site, this analysis is limited in its temporal complexity. Current landscape pattern is not always reflective of current population pattern due to the delayed response of plants to landscape change. The temporal scale of landscape change is different from the temporal scale of flora species' response to landscape change such that present species occurrences can still reflect past landscape patterns (Cousins et al., 2007; Helm et al., 2006). Remnant patches can exist within a population because there is a considerable time span between migration of new individuals to a patch and the final extinction of species in a patch during succession (Eriksson, 1996).

Therefore, it can be more useful for explaining population pattern to compare past landscapes with present landscapes to see how it has changed.

The results of this study are limited to the scale and the study site at which they were conducted. However, the methods are applicable elsewhere. Significant findings at one spatial scale will change as the scale changes because the detection of spatial and temporal variation depends on the scale at which it is sampled and the size of the mapping unit (Gustafson, 1998). As well, site specific management strategies will affect the spatial distribution and pattern of sand patches such that variables which significantly affect a population at one site may not be present or significant at another site due to the differences in the function and composition between the different landscapes.

4.3 Future Research

Future research could be done to improve the predictive capability of the logistic regression model and increase the accuracy of habitat suitability predictions for hairy prairie-clover. Variables such as elevation, slope angle, slope aspect, amount of incoming solar radiation, surface soil moisture, and surface wind speed have been identified as key ecological factors affecting population viability in plant species (Wolken et al., 2001; Parolo et al., 2008). These variables can be derived from remote sensing imagery such as Lidar, which is capable of deriving a high accuracy digital elevation model. For example, information on surface wind speed and amount of incoming solar radiation could be useful in explaining the localized distribution of hairy prairie-clover on a sand patch. It could give such information as locations where wind speed is low enough for seed deposition to occur, or surface temperature which can affect seed germination (Schellenberg et al., 2010), respectively. Further, demographic information on hairy prairie-clover, such as seed dispersal pathways and distances, could help improve estimations of sand patch isolation. Adding these ecological variables to the model could further explain hairy prairie-clover presence/absence on sandy areas.

Future research should also include the replication of time and space to compare and validate the results of this study. Ecologically meaningful relationships found in this study can be verified by comparing the same landscape over time and/or comparing different landscapes within the same time.

This research also provides a base for future analysis of landscape function and composition as represented in Noss's (1990) hierarchical structure of ecosystem studies. While

this research found that structure played a significant role in the distribution of hairy prairie-clover on the landscape, about 71% of the variation in patch occupancy was left unexplained. Future research considering the functional and compositional components of the landscape may further help explain hairy prairie-clover distribution. For example, landscape function may be significant in explaining hairy prairie-clover distribution because it has been found that fire plays a key role in maintaining the density and occurrence of legumes on the prairies (Leach and Givnish, 1996).

4.4 References

Cousins, S., Ohlson, H., Eriksson, O., 2007. Effects of historical and present fragmentation on plant species diversity in semi-natural grasslands in Swedish rural landscapes. *Landscape Ecology* 22: 723-730.

Elith, J., Graham, C., Anderson, R., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R., Huettmann, F., Leathwick, J., Lehmann, A., Li, J., Lohmann, L., Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J., Peterson, A., Phillips, S., Richardson, K., Scachetti-Pereira, R., Schapire, R., Soberon, J., Williams, S., Wisz, M., Zimmermann, N., 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29: 129-151.

Eriksson, O., 1996. Regional dynamics of plants: a review of evidence for remnant, source-sink and metapopulations. *Oikos* 77: 248-258.

Fahrig, L. and Merriam, G., 1994. Conservation of fragmented populations. *Conservation Biology* 8: 50-59.

Freckleton, R. and Watkinson, A., 2002. Large-scale spatial dynamics of plants: metapopulations, regional ensembles and patchy populations. *Journal of Ecology* 90: 419-434.

Godwin, B. and Thorpe, J., 2007. Targeted surveys for plant species at risk in Elbow, Dundurn and Rudy-Rosedale AAFC-PFRA Pastures, 2006. Agriculture and Agri-Food Canada-Prairie Farm Rehabilitation Administration. *Saskatchewan Research Council Pub. No. 11997-1E07.*

Geerken, R., Zaitchik, B., Evans, J., 2005. Classifying rangeland vegetation type and coverage from NDVI time series using Fourier filtered cycle similarity. *International Journal of Remote Sensing* 26: 5535-5554.

Gustafson, E., 1998. Quantifying landscape spatial pattern: What is the state of the art? *Ecosystems* 1: 143-156.

He, Y., Guo, X., Cheng Si, B., 2007. Detecting grassland spatial variation by a wavelet approach. *International Journal of Remote Sensing* 28: 1527-1545.

Helm, A., Hanski, I., Partel, M., 2006. Slow response of plant species richness to habitat loss and fragmentation. *Ecology Letters* 9: 72-77.

Herrera, L., Laterra, P., Maceira, N., Zelaya, K., Martinez, G., 2009. Fragmentation status of tall tussock grassland relicts in the flooding Pampa, Argentina. *Rangeland Ecology and Management* 62: 73-82.

Hugenholtz, C. and Wolfe, S., 2005. Recent stabilization of active sand dunes on the Canadian prairies and relation to recent climate variations. *Geomorphology* 68: 131-147.

Leach, M. and Givnish, T., 1996. Ecological determinants of species loss in remnant prairies. *Science* 273: 1555-1558.

Noss, R., 1990. Indicators for monitoring biodiversity: A hierarchical approach. *Conservation Biology* 4: 355-364.

Parolo, G., Rossi, G., Ferrarini, A., 2008. Toward improved species niche modelling: Arnica Montana in the Alps as a case study. *Journal of Applied Ecology* 45: 1410-1418.

Schellenberg, M., Henderson, D., Bolton, J., St. Pierre, R., 2010. Do the germination temperature characteristics of the species at risk hairy prairie-clover (Dalea villosa) differ from the more common prairie clovers (D. purpurea and candida) found in the Canadian Prairies? Poster presentation at the 9^{th} *Prairie Conservation and Endangered Species Conference*. URL: http://www.pcesc.ca/postercat5.html [Last accessed: Dec. 9, 2010].

Vitt, P., Havens, K., Kendall, B., Knight, T., 2009. Effects of community-level grassland

management on the non-target rare annual *Agalinis auriculata*. *Biological Conservation* 142: 798-805.
Walker, B., 1995. Conserving biological diversity through ecosystem resilience. *Conservation Biology* 9: 747-752.
Williams, N., Morgan, J., McCarthy, M., McDonnell, M., 2006. Local extinction of grassland plants: The landscape matrix is more important than patch attributes. *Ecology* 87: 3000-3006.
Wolken, P., Sieg, C., Williams, S., 2001. Quantifying suitable habitat of the threatened western prairie fringed orchid. *Journal of Range Management* 54: 611-616.
Zhang, C. and Guo, X., 2007. Measuring biological heterogeneity in the northern mixed prairie: a remote sensing approach. *Canadian Geographer* 51: 462-474.